WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

	(51) International Patent Classification 4:	:		(1)	1) International Publication Number:	WO 89/ 06286	
	C12Q 1/68, C07H 21/00 C12P 21/02, C12N 15/00 G01N 33/68, A61K 45/00		A2		3) International Publication Date:	13 July 1989 (13.07.89)	
i	(21) International Application Number:	PCT/US	88/045	504	(74) Agent: FREEMAN, John, W.;	Fish & Richardson, One	

(22) International Filing Date: 16 December 1988 (16.12.88)

(31) Priority Application Number: 136,618

(32) Priority Date: 22 December 1987 (22.12.87)

(33) Priority Country:

(71) Applicant: THE CHILDREN'S MEDICAL CENTER CORPORATION [US/US]; 300 Longwood Avenue, Boston, MA 02115 (US).

(72) Inventors: KUNKEL, Louis, M.; 5 Braeburn Road, Hyde Park, MA 02136 (US). MONACO, Anthony; 51 Queenberry Street, #1, Boston, MA 02115 (US). HOFFMAN, Eric, P.; 133 Waban Street, Newton, MA 02158 (US). KOENIG, Michel; 51 Park Drive, No. 9, Boston, MA 02115 (US).

Financial Center, Suite 2500, Boston, MA 02110-2658

(81) Designated States: AT (European patent), AU, BE (European patent), BR, CH (European patent), AC, BE (European patent), BR, CH (European patent), DE (European patent), JK, FI, FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), NO, SE (European patent), SU.

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: PROBES FOR AND METHODS OF DIAGNOSIS FOR MUSCULAR DYSTROPHY

(57) Abstract

\$ 15

The invention relates to a muscular dystrophy (MD) probe comprising a substantially purified single-stranded nucleic acid sequence capable of hybridizing to a region of DNA on a human X chromosome between the deletion break point at Xp21.3 and the translocation break point at X;11. The invention also relates to a 14kb cDNA corresponding to the complete MD gene and probes produced therefrom useful in genetic methods of diagnosis of MD. Furthermore, the invention relates to the polypeptide, dystrophin, which corresponds to the MD gene product, and antibodies thereto that are useful in a variety of methods for immunodiagnosis of MD.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

	ΑT	Austria	FR	France	ML	Mali
	AU	Anstralia	GA	Gabon	MR	Mauritania
	BB	Barbados	GB	United Kingdom	MW	Malawi
	BE	Belgium	HU	Hungary	NL	Netherlands
	BG	Bulgaria	I	Italy	NO	Norway
	BJ	Benin	JP	Japan	RO	Romania
	BR	Brazil	KP	Democratic People's Republic	SD	Sudan
	CF	Central African Republic	_	of Korea	SE	Sweden
	CG	Congo	KR	Republic of Korea	SN	Senegal
	Œ	Switzerland	ū	Liechtenstein	SU	Soviet Union
	CM	Cameroon	LK	Sri Lanka	TD	Chad
•	DE	Germany, Federal Republic of	LU	Luxembourg	TG	Togo
	DK	Denmark	MC	Monaco	ŪS	United States of America
	DA ET	Waland	MC	Madagagese		Ommi 0200 21.20.

PROBES FOR AND METHODS OF DIAGNOSIS FOR MUSCULAR DYSTROPHY

Background of the Invention

This application is a continuation-in-part of U.S. Patent Application Serial No. 890,694 which is incorporated herein by reference.

The work reported herein was supported by grants from the United States government including N.I.H. grants HD 18658 and NS23740. The government has 10 certain rights in this invention.

This invention relates to the detection and treatment of hereditary disease, and in particular to the detection and diagnosis of Duchenne, Becker and Outlier muscular dystrophies (MD) by various methods.

Duchenne muscular dystrophy (DMD) is an X-linked recessive genetic disorder which affects about 1 in 3,300 males. Traits associated with DMD (DMD phenotype) are well known and may include elevated creatine phosphokinase levels in serum (at least 10x the normal level), delayed development of motor function, and muscle weakness characterized by the replacement of muscle fiber with adipose and fibrose tissue accompanied by a marked variation in muscle size. Until recently, carrier identification in DMD families generally was accomplished by detecting elevated levels of creatine phosphokinase in serum.

Becker muscular dystrophy (BMD) is also an X-linked recessive genetic disorder, but occurs at only 10% of the frequency of DMD. BMD is a more benign form of muscular dystrophy which follows a less rapid

clinical course than DMD. Both DMD and BMD are caused by mutations in the DMD gene located in the Xp21 region of the short arm of the human X chromosome.

Outlier muscular dystrophy (OMD) is a mild 5 Duchenne/severe Becker disorder. It forms a subgroup that comprises approximately 10 percent of individuals inflicted with DMD.

Restriction fragment length polymorphisms
(RFLPs) occur when the genomic DNA sequence in a

10 population of animals varies such that complete
digestion with a restriction endonuclease will result in
a set of restriction fragments of one or more different
size arrays. The differences are due to the presence or
absence in the DNA sequence of sites recognized by the

15 endonuclease; a sequence in which the site is present
will be cut by the endonuclease giving one array, and a
sequence in which the site is not present will fail to
be cut by the endonuclease, giving another array.

It has been shown that RFLPs which are located 20 near a gene associated with a hereditary disease can be used to determine the likelihood of a particular person having a defective gene and thus having the disease. Bakker et al., The Lancet, March 23, 1985, at 655 describes "eleven RFLP-markers... on the short arm of 25 the X chromosome [that] are useful in the diagnosis of DMD since they bridge the Duchenne locus at genetic distances varying between 3 and 20 centiMorgans."

Summary of the Invention

Applicants have cloned 14 kb of human cDNA 30 substantially corresponding to the MD gene, and parts of the surrounding genomic DNA. Applicants have also isolated and characterized cDNA corresponding to the homologous locus in the mouse. Dystrophin, the polypeptide product of the human MD locus, and its mouse

homologue (mMD), has been identified using antibodies directed against fusion polypeptides containing two distinct regions of MD and mMD cDNA. Parts of the complete MD cDNA and surrounding genomic DNA were also used to define RFLPs in the Xp21 region of the short arm of the X chromosome that are useful in diagnosis of MD mutations.

In a first aspect, the invention features an MD probe having a substantially purified single-stranded nucleic acid sequence capable of selectively hybridizing to a region of DNA on a human X chromosome between the deletion break point at Xp21.3 (defined below) and the translocation break point at X;11 (defined below).

"Selectively hybridizing" means hybridizing to the desired sequence under conditions that are sufficiently stringent to generally prevent hybridization to other sequences encountered in the sample. "Substantially purified" or "substantially pure" means a substance that is sufficiently separated from other elements of the environment in which it naturally occurs to function in assays according to the invention.

In preferred embodiments, the nucleic acid sequence is incapable of selectively hybridizing to other parts of the short arm of the X chromosome, i.e., the region of DNA between the deletion break point at Xp21.3 and the p terminus of the X chromosome, and to the region of DNA between the translocation break point at X;11 and the centromere of the X chromosome.

In an additional preferred embodiment, the MD 30 probe has a nucleic acid sequence of at least 10 base pairs in length from a portion of the sequence of any one of the six cDNA probes, that together span the 14kb cDNA corresponding to the human MD gene, deposited on July 15, 1987 with the A.T.C.C. under the twelve access

numbers 57666 to 57677, inclusive. Each probe is deposited in a vector in $\underline{\mathbf{E}}$. coli and in the form of a purified plasmid. Hence, each probe is available under two access numbers.

In a second aspect, the invention features an MD probe having a substantially purified single-stranded nucleic acid sequence capable of selectively hybridizing to at least a part of the human MD gene. The human MD gene, as used herein, is that portion of human DNA on the short arm of the X chromosome in which a mutation can give rise to an MD phenotype. Exons in this DNA are substantially represented by the 14 kb cDNA deposit discussed below. The sequence for most of this cDNA is depicted in Fig. 5. The invention also includes probes directed to or derived from genomic DNA, including intronic sequences in the MD gene.

In preferred embodiments, the nucleic acid probe sequence is incapable of selectively hybridizing to regions of DNA outside the MD gene.

The probe of the invention can be used for analyzing a sample of genomic DNA of a human subject for the presence of a mutant MD gene.

In a third aspect, the invention features a method for diagnosing muscular dystrophy in human patients involving contacting the probes outlined above with a sample and detecting whether these probes hybridize to nucleic acid in the sample.

In preferred embodiments, a series of probes is used in the method of diagnosis, in which the series, in combination, spans a 14 kb sequence of MD DNA or RNA, encompassing an MD gene. For example, probes according to the invention are used for detecting RFLPs that are closely associated (linked) with mutations in the MD gene causing an MD phenotype. "Closely associated"

means that the likelihood of a genetic recombination event occurring between the sites of an RFLP and a MD mutation is very low (less than 5%). The advantage of such probes is that one can use them to predict, with a high degree of confidence, the probability that an individual (within a family with a history of MD) having a particular RFLP phenotype will also have an MD mutation.

In a fourth aspect, the invention features a

10 substantially pure mammalian polypeptide, dystrophin,
encoded by an MD gene. This polypeptide has both human
and murine forms. The invention also includes
substantially pure polypeptides that are immunologically
cross-reactive with dystrophin, or which have

15 substantially the same amino acid sequence as a 15 amino
acid sequence of dystrophin. By "polypeptide" we mean
to include large polypeptides and proteins.

"Substantially the same" means that the amino acids in
such a sequence may be substituted by other amino acids

20 having equivalent side groups, i.e., conservative
substitutions are permissible.

Two polypeptides are immunologically cross-reactive if they are bound to the same specific antibody or family of specific antibodies. Particularly useful types of cross-reactive polypeptides are those that can be used to challenge an animal to produce an antibody that reacts with dystrophin. The invention also includes polypeptides that are useful for competitive assays, i.e., polypeptides that compete with dystrophin for binding sites on specific antibodies.

In a fifth aspect, the invention features substantially purified nucleic acid encoding dystrophin, or polypeptides immunologically cross-reactive with dystrophin.

A preferred embodiment of this nucleic acid comprises an approximately 14 kb pair sequence of cDNA or RNA corresponding to a nucleic acid sequence of a MD gene, or fragments thereof greater than 45 bases. This nucleic acid may correspond to a human MD gene or a murine Dmd gene.

In a sixth aspect, the invention features antibodies to dystrophin or polypeptides that are immunologically cross-reactive with dystrophin.

The invention also includes the use of these antibodies in a method of immunodiagnosis for MD involving contacting a sample of tissue (particularly muscle tissue) with one of these antibodies and detecting the presence of antibody-dystrophin

15 complexes. Complexes that indicate the presence of dystrophin of normal size and abundance in such tissue

dystrophin of normal size and abundance in such tissue generally indicate the absence of MD. This aspect of the invention includes the numerous well-known formats of immunodiagnostics, such as ELISA, and Western blots.

In one embodiment, this method is performed by comparing the molecular weight (M.W.) of dystrophin, or fragments thereof, that are detected in the complex to the molecular weight of a normal control dystrophin polypeptide. In this method, the presence of fragments

25 of dystrophin, i.e., dystrophin of abnormal size, or the

complete absence of detectable dystrophin, indicates MD. A normal control is established by repeated analyses of the dystrophin polypeptide found in normal individuals without MD or individuals without MD related dystrophins.

In general, the normal dystrophin polypeptide is approximately 400,000 daltons in size and is absent or undetectable in DMD patients. However, BMD patients

WO 89/06286 PCT/US88/04504

and a small percentage of patients diagnosed as those with DMD exhibit dystrophin of an abnormal size, i.e., in the form of smaller M.W. fragments.

In a preferred embodiment, this method features 5 the additional step of comparing the molecular weight of the sample antigenic polypeptide to the molecular weight of BMD and DMD controls. In this method diagnosis of DMD is made in the presence of a pattern comparable to the DMD control, diagnosis of BMD is made in the 10 presence of a pattern comparable to the BMD control, and the absence of MD is indicated by a pattern comparable to the normal control. The BMD and DMD controls are established by repeated assays of dystrophin in muscle samples in patients diagnosed as having BMD or DMD. 15 molecular weight of the dystrophin or fragments thereof are determined for each type of MD and are recorded and averaged to create a control pattern. A pattern "comparable" to these controls is a pattern that is identical to, or sufficiently similar to, these controls 20 to allow a diagnosis of one type of MD, i.e., either BMD or DMD.

The invention also features a method of therapy for MD involving the insertion of cDNA, or a fragment thereof, corresponding to the MD gene into a vector and reintroducing these genetically altered cells back into the patient. The cells are injected into the bloodstream or muscle tissue to produce dystrophin in an amount effective to control the degeneration of muscle fibers and to control the proliferation of connective tissue within muscle fibers.

The methods of the invention can be used for prenatal (fetal) screening as well as in the detection of heterozygotic individuals carrying MD gene mutations but exhibiting no symptoms.

20

25

30

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

<u>Description of the Preferred Embodiments</u>

The Figures will first briefly be described.

Drawings

Fig. 1 is a standard ISCN (1981) ideogram of the p terminal end (pter) of the human X chromosome showing the approximate positions of cloned DNA fragments and specific deletion and translocation break points;

Fig. 2 is a diagrammatic representation of 220kb of cloned DNA from the human Xp21 chromosome, showing regions deleted in some DMD human males;

Fig. 3 is the DNA sequence of two parts of the MD gene used as probes of the invention;

Fig. 4 is a diagrammatic representation of 220kb of cloned DNA showing the location and size of cloned DNA isolated from librares of human chromosomal DNA;

Fig. 5 shows 13 kb of the 14 kb nucleotide sequence corresponding to the human MD gene.

Fig. 6A is a restriction map of the 14 kb cDNA encoding the MD gene;

Fig. 6B is a diagrammatic representation of cDNA fragments that in combination span the entire 14 kb cDNA;

Figs. 7A and 7B are partial nucleotide and predicted amino acid sequences of human and murine MD cDNAs;

Fig. 8 is the amino acid sequence of the dystrophin polypeptide.

Fig. 9 is a photograph of a Southern blot analysis of genomic DNA isolated from a DMD family,

probed with pERT87-8, and a schematic representation of the genotype of each family member, circles represent females, squares represent males, darkened areas indicate the presence of a MD mutant gene.

Fig. 10 is a schematic of a DNA family polymorphism study (for Family A) in which each vertical bar represents an Xp21-p22 chromosome region, and each letter designates a different polymorphic DNA segment;

Figs. 11A and 11B are photographs of 10 immunostained Western blot gels containing muscle samples of DMD, BMD and OMD patients.

Fig. 12 is a photograph of an immunostained Western blot containing muscle samples of BMD patients and normal non-MD controls.

15 Fig. 13 is a set of charts depicting the distribution of dystrophin biochemical phenotypes.

Human MD Probes and Nucleic Acid

Probes of the invention contain nucleic acid homologous to DNA of the MD gene or to DNA from a region of DNA close to the MD gene. A description of the DNA region containing segments to which the probes are homologous follows.

Referring to Fig. 1, the approximate positions of previously cloned regions of the human X chromosome 25 (D2, RC8, 99.6, 754, OTC C7, B24, and L128) are shown relative to a standard (ISCN 1981) ideogram of the short arm (p) of the human X chromosome. Also shown is the region of DNA deletion in the X chromosome of a male patient (BB) suffering from three X-linked disorders, including DMD; the deletion is described by Francke et al., Am. J. Hum. Genet. 37:250 (1985), and the cell line is available from Human Genetic Mutant Cell Repository

in Camden, N.J. (cell line repository number GM7947A). The deletion is defined by two break points in the DNA, Xp21.3 and Xp21.1. Also marked is a break point X;11 (p. 21.1; a 13.5), which is attributable to the balanced reciprocal translocation described by Greenstein et al., Cytog. Cell Genet. 27:268 (1980) (cell line available from Human Genetic Mutant Cell Repository in Camden, N.J., cell line repository identification number GM1695).

Break points Xp21.3 and X;11 approximately

define the boundaries of the DMD gene, which lies within
the Xp21.2 chromosome band region. More specifically,
the MD gene extends approximately 1x10⁶ bp beyond both
the BB deletion breakpoint towards the p terminus and
beyond the X;11 translocation breakpoint toward the

centromere.

Following are examples of probes and nucleic acids suitable in this invention. These examples are not limiting to the invention and those skilled in the art will recognize other methods suitable to obtain such probes. In particular, the probes described below can be used to probe other libraries of other animals to isolate equivalent DNA or RNA.

Example 1: pERT87, A Genomic Clone

In order to isolate the probes of the

invention, a library of DNA from around the MD gene area
was constructed by enriching for this DNA using
differential hybridization. The procedure is described
in detail by Kunkel et al., PNAS 87:4778 (1985).
Accordingly, 250 μg of DNA from the male patient BB

was isolated, sheared by sonication to a mean size of
1,000 base pairs (bp), and hybridized to 1.25 μg of
Mbo-1 cleaved DNA from a 49, XXXXX lymphoid cell line
(GM1202), as described by Palmer et al., Cell 37:171

(1984). The DNAs were then heated to 100°C for 5 minutes, cooled on ice, and added to a final reaction volume of 2.5 ml of 7% phenol. The mixture was shaken intermittently for 37 h over a 5 day period (Kohne et al., Biochemistry 16:5329 (1977)). After chloroform extraction and dialysis to remove salts, the hybridized DNA was ethanol precipitated. These conditions allow Mbo-1 fragments from the 49, XXXXY DNA, with no complementary sequence in the deleted DNA region of the BB male, to self-hybridize and thus be cloned as Mbo-1 fragments. 5 μg of the precipitated DNA was ligated to 0.1 μg BamH1-cleaved dephosphorylated pBR322 and transformed into E. coli strain MC1061. (The BamH1 ends are compatible in ligation with the Mbo-1 ends of the chromosomal DNA.) 3,000 colonies were isolated.

As described by Kunkel et al., supra, the isolated colonies were screened using standard techniques to determine which one contained DNA that hybridized with DNA from normal individuals but not with 20 DNA from the male BB. One clone, pERT87, hybridized to a 1.1 kb HindIII fragment of normal human DNA, but not to BB DNA. The intensity of hybridization was dependent upon the X chromosome content of the DNA being probed. Hybridization was also detected in a rodent-human hybrid 25 cell line having an intact human X chromosome, but not with a cell line having DNA only from the human chromosome region Xpl1.3 to Xqter. Thus, the cloned region in pERT87 is on the short arm (p) of the X Three other clones gave similar results, 30 although each clone hybridizes to a different HindIII fragment.

The clones were analyzed for their ability to hybridize with DNA from a variety of cell lines having

deletions or translocations of the human X chromosome.

pERT87 was localized to DNA within the Xp21-Xp22

region. The location of pERT87 was determined by its
lack of hybridization to the region Xp21-Xter (using
cell line X;11) and to the region p21-pter (using a cell
line with a break point at Xp21.3, Fryns et al., Clin.,
Genet. 22:76, (1982)), and by its hybridization to the
region Xp22-Xqter (using a cell line described by
Mahandas et al., PNAS 76:5779 (1979)).

10 As described in Monaco et al., Nature 316:842
(1985), pERT87 was used to probe the DNA of 57 unrelated
DMD males. pERT87 did not hybridize to DNA from five of
the males indicating that they had deletions of DNA
around the cloned DNA in pERT87. Assuming that the DMD
15 phenotype in these five males results from deletions
within the MD gene, this result provides evidence that
pERT87 contains DNA homologous to a region of DNA within
the MD gene.

Example 2: pERT87 Derivatives

- In order to isolate more DNA from the MD region, two human genomic libraries were screened with pERT87 by the procedure described in Monaco, supra; one was constructed in Charon 35 (Loenan et al., Gene 26:171 (1983)) and one in EMBL-3 (Frischauf et al., J. Molec.
- 25 Biol. 170:827 (1983). Using standard techniques, chromosomal "walks" were performed in both directions from the region homologous to pERT87 using pERT87 as a probe for the above libraries. Five bacteriophage clones, including W1 and W2-R in Fig. 4, were isolated,
- 30 and small unique-sequence subclones, in pBR322, pUC18 or Blue-Scribe (Stratagene), were constructed. These subclones were restriction mapped and small fragments free of human repetitive elements identified. Unique

sequence subclones near the extreme ends of the human inserts in these clones were used as probes for the next walking step in the same libraries. Three of these subclones, pERT87-18, -8, and -1 contained DNA absent from the above five DMD boys with genomic deletions. These subclones span an area of 38 kb. Further chromosome walks in the EMBL-3 library extended the cloned area to 220 kb (shown in Figs. 2 and 4). These clones include pERT87-15, -14, and -27.

10 To determine the positions of these cloned regions relative to the MD gene, chromosomal DNA was isolated from 53 boys with DMD or the related disease, Becker muscular dystrophy (BMD) -- the gene for which is also located within band Xp21--who have deleted areas of 15 DNA in their X chromosome. This DNA was tested, using standard Southern blot analysis (see below), against the subclones in order to determine the break points of the DNA deletions in the 53 boys relative to the cloned This was simply done by first ordering the 20 subclones as shown in Fig. 4, and then determining which subclones contained DNA present in the deleted chromosomes (thus being outside the break point) and those which contained DNA not present in the deleted chromosome (thus having DNA within the break point). 25 Subclones that are positioned between these two types of subclones and have DNA present in the deleted chromosomes, must have DNA in which the deletion break point occurred. Once localized to a region within the · cloned DNA, the break points were further localized 30 within 100-1,000 bp using restriction enzyme analysis. The results of these experiments are summarized in Fig. 2.

Fig. 2 depicts a schematic drawing of 220 kb of contiquous human genomic DNA with a kb scale.

The darkened blocks represent regions of repeated sequences, cross hatched boxes represent moderately repeated sequences, and open boxes represent unique sequences. Numbers below these blocks represent pERT87 subclones. Line sequences are represented by an L, as described in Singer, Cell 28:433 (1981). As shown in Fig. 2, fourteen deletions were found to break at the left side of the cloned DNA and extend in a rightward direction (see arrows) towards the centromere. Nine 10 deletions were found that extend in the opposite direction. The majority of the deletions have independent break points, but for simplicity the deletions have been schematically presented together in Fig. 2. Since the deletion breaks, which give rise to a 15 DMD or BMD phenotype, occur within the pERT87 region and extend in both directions, this analysis demonstrates that pERT87 and the subclones are clearly within the MD gene.

The profile in Fig. 2 indicates that the MD

20 gene is located on a large segment of DNA. The fact
that large deletions resulting from mutations in the
gene yield a DMD phenotype similar to that found in
other DMD boys who do not have deletions indicates that
the gene product can be either completely missing or
25 aberrant yet still yield a similar clinical picture.
The gene responsible lies on a very large segment of X
chromosome DNA, and is estimated to be approximately
2,000 kb (Fig. 1).

Referring to Fig. 3, the DNA sequence of 30 several subclones are determined by standard procedures, using dideoxy analysis. Representative sequences from pERT87-4 (upper sequence) and -25 (lower sequence) are presented. Probes suitable for the invention can

readily be synthesized from any parts of these sequences, or any other parts of the 14 kb cDNA sequence corresponding to an MD gene described below. Such probes can also be used for further chromosome walking.

Example 3: cDNA

Standard Northern analysis of RNA from a variety of human fetal and adult tissues revealed a single mRNA estimated to be 16 kb from the total RNA or polyA-selected RNA of fetal skeletal muscle which hybridizes to pERT87-25. This approximately 16 kb 10 transcript is the mRNA coding for the product of the MD gene. It is possible to isolate this mRNA, using standard hybridization techniques, and the probes of the invention, and to prepare a cDNA library from it. clones in this cDNA library are suitable as probes of 15 the invention, and for isolating other clones by chromosome walking in human genomic DNA libraries by standard techniques towards intronic DNA in the MD gene and to the X;11 and Xp21.3 break points.

For example, by using standard techniques, an 20 oligo (dT) primed cDNA library is constructed in the phage vector lambda gtll using a poly A-selected RNA sample from human fetal skeletal muscle. Small 5 -ended clones were isolated from this library. Similarly, a human fetal skeletal muscle cDNA library can be 25 constructed in the phage vector lambda gt10 using standard techniques as described in Koenig et al., Cell, 50:509-517 (1987).

Such a library was used to isolate six cDNA inserts detected with the above 5'-end probes. All six 30 ended within an estimated 100 bp interval (the five longest within a 50 bp interval). This result indicates that the 5' end of the transcript is probably present in

these clones. Two additional "walks" taken in the 3' direction using the most 3' segment clones from these probings also ended within 100 bp of each other, indicating that the 3' end of the transcript is probably present in these clones. As a result, the actual size of the mRNA or cDNA corresponding to the human MD gene is known to be 14 kb rather than 16 kb estimated in earlier studies described above.

5

Once the 5' most and 3' most ends of the human

MD gene transcript are determined, the entire 14 kb cDNA
is sequenced using standard techniques as described in
Koenig et al., Cell, 50:509-517 (1987). The nearly
complete sequence is described in Fig. 5. The actual
cDNA can be obtained from the series of cDNA probes

deposited with the ATCC that in combination span the
entire human MD gene. These probes are deposited under
the twelve access numbers 57666 to 57677 and are
depicted, in order, next to the 14 kb MD gene in Figure
6.

These cDNA clones are tested to determine whether any clones are particularly suitable as probes of the invention. The clones are hybridized, using standard methods, to cell lines having the Xp21.3 and X;11 mutations and to mRNA from human cell lines.

Suitable probes are homologous to DNA from within the Xp21.3 -X;11 region or within the mRNA. Referring to Fig. 1, suitable probes have homology to normal human cell lines but at least part of their DNA will not have homology to DNA from cell line GM7947A and the

nontranslocated DNA of the cell line GM1695. A further test entails a cytological study to determine whether the clone hybridizes to the Xp21.2 chromosomal band, and thus would be suitable.

Detection of RFLPs

The probes of the invention are useful in detecting RFLPs that are closely linked with a mutated MD gene that gives rise to the DMD phenotype.

Detection of such RFLPs allows diagnosis both of women who are carriers of a mutated MD gene and of unborn children who have a mutated MD gene.

Example 5: Using pERT87 Derivatives

Using standard techniques, the subclones

10 pERT87-8, and pERT87-15 were tested for their ability to
detect human DNA RFLPs by hybridizing to nitrocellulose
filters containing immobilized DNA samples from four
unrelated females that had been cleaved with 24
different restriction enzymes. The frequencies of the

15 RFLPs found in the human population were estimated by
examining DNA obtained from 37 unrelated persons. The
results of the search are displayed in Table 1, below.
Seven potential polymorphic restriction enzyme sites
were identified as individual variations in the pattern
of restriction fragments observed. Two of the
subclones, pERT87-1 and pERT87-8, each detected putative
RFLP's of two different enzymes, while one probe,
pERT87-15, detected RFLPs of three different enzymes.

Table 1

25

			Allele <u>Sizes (kb)</u>		Allele Frequency	
	1. Subclone	<u>Enzyme</u>				
			p*	d **	p	P
	2. pERT87-1	BstNI	3.1	2.5/0.6	0.63	0.37
30	3.	XmnI	8.7	7.5	0.66	0.34
	4. pERT87-8	BstXI	4.4.	2.2	0.6	0.4
	5.	TagI	2.7/1.1	3.8	0.71	0.29

- 18 -

	. Subclone Enzyme		A	llele	Allele	
1.			Sizes (kb)		Frequency	
6.	pERT87-15	BamHI	7.1/2.3	9.4	0.62	0.38
7.		TaqI	3.1	3.3	0.67	0.33
8.		XmnI	1.6/1.2	2.8	0.68	0.32

5 * common allele

** rare allele

The next step was to screen the RFLPs clones for linkage to the DMD phenotype. An RFLP is linked to the MD gene within a particular pedigree (family) if within the pedigree, persons whose DNA contains a mutant MD gene giving rise to the DMD phenotype have a polymorphic distribution different from those persons within the pedigree whose DNA does not contain a mutant MD gene. The RFLP screening was performed on pedigrees known to harbor a mutant MD gene.

The first step was to obtain DNA from a subject. In the method described below, peripheral blood lymphocytes are used; other possible sources of DNA are amniotic fluid, chorionic villus, and fetal trophoblasts.

Using standard techniques, genomic DNA is isolated from cell nuclei of whole blood leukocytes as described by Aldridge et al., Am. J. Hum. Gen., 36: 546 (1984). This genomic DNA is then analyzed to determine whether linkages exist between the RFLPs of Table 1 and defective MD genes of certain families.

An example of such linkage is presented in Figure 9. In the example, DNA was isolated from individuals within a family in which some people had a mutated MD gene (indicated by darkened symbols) the DNA was cleaved with BstXI, and probed with pERT87-8.

Electrophoresis, blotting, and hybridization were as described above. On BstXI digestion of genomic DNA, pERT87-8 hybridizes to 4.4 and 2.2 kb BstXI fragments in females heterozygous for the RFLP. The carrier mother exhibits the heterozygous pattern of the 4.4/2.2 kb BstXI alleles, as do her two carrier daughters (lanes 2,4). The lower allele (2.2 kb) is present in both affected DMD sons (lanes 5,6) while only the upper allele (4.4 kb) is present in the unaffected son 10 (lane 3). The father of this family has a 4.4 kb BstXI hybridizing fragment (data not shown). The female individuals in lanes 2 and 4 were suspected of being carriers of the DMD trait based on elevated CPK values. Fragment sizes were calculated by comparison with 15 32 P-end-labelled HindIII-cleaved λ phage DNA. Thus, the mutated MD gene is linked to the BstXI RFLP in this family, and individuals with 2.2 kb fragment also have a mutated MD gene.

Linkage between each RFLP listed in Table 1 and a defective MD gene does not occur in every pedigree. Accordingly, a combination of probes, detecting a plurality of RFLPs, is preferably used when examining DNA of an individual. Examining a combination of RFLPs listed in lines 2, 3, 5, 6, and 7 of Table 1 with the appropriate probes was found to be informative in 25 of 28 (89%) pedigrees in which members were known to have a mutated MD gene. It is also possible to use probes of the invention for presymptomatic screening of humans without a DMD phenotype for deletions (of greater than 100 bps) in the MD gene.

Example 6: Using cDNA Probes

Although RFLP DNA marker studies are usually very useful, they are often impossible or inconclusive because family members may be unavailable, new mutation

events cannot be detected reliably, meiotic crossovers within the gene occur frequently and the inheritance pattern (X-linked versus autosomal) cannot be determined with certainty in sporadic cases.

Detection of molecular deletions with the complete gene cDNA probes of the invention eliminates most of these difficulties. For example, the probes of the invention deposited with the ATCC under the access numbers 57666 to 57677, inclusive, are very useful 10 diagnostic tools as described below.

Nuclear DNA of patients was isolated from blood leucocytes, lymphoblastoid cell lines and chorionic villi, cleaved with restriction endonucleases, transferred to Hybond (Amersham) filters, and hybridized 15 with oligo-labelled probes, by standard techniques. cDNA segments covering the complete MD gene in a 5' to 3' direction, cloned in M13 vectors Bluescribe or Bluescript (Stratagene), are used as probes. example, Figure 6 shows the probes that have been 20 deposited with the A.T.C.C. Probes 1-2a and 2b-3 comprise the most 5' 2.6 kb BqlII-EcoRI fragment. Probe 4-5a is a 1.8 kb fragment and probe 5b-7 has a 2.6 kb insert. Probe 8 is a 0.9 kb EcoRI fragment. Probe 9-14 is a 6.1 kb fragment that includes the 3' exons and the 25 3' untranslated region.

Deletions in the MD gene do not occur in a random distribution along the cDNA, but a large percentage of the deletions can be detected with only a few probes. In a total of 104 DNA samples of unrelated DMD boys 53 exhibited a deletion of part of the MD gene. For diagnostic purposes, 27 deletions are detected with probe 8, and 15 others with probe 1-2a. These two probes detect nearly 80% of the deletions observed (42 of 53).

Particularly intriguing is the large number of deletion breakpoints found to originate within a single intron defined by probe 5b-7. One intron defined by the exons contained in this probe exhibits 19 deletion

5 breakpoints (36% of the 53 total deletions found).

Clearly this particular intron contains a characteristic that imparts a propensity for initiating deletion events. The multiple breaks could be due to a sequence-specific rearrangement hot spot, or an extraordinarily large genomic distance spanned by this intron.

The following cost-efficient strategy for the DNA-diagnostic work-up of DMD/BMD families is based on these results. Initially, DNA from affected 15 individuals, their mothers, obligate carriers and normal male and female controls is cleaved with Bq1II and HindIII, electrophoresed and blotted onto nylon membranes that can be easily rehybridized. filters are first hybridized with probe 8, exposed to 20 X-ray film, and then rehybridized with probes 1-2a and 2b-3 in a sequential fashion. If no deletion is detected, probes 5b-7 and 4-5a are used next, and probe 9-14 is used in the final round. It is advisable to use two different restriction enzymes because fragments of 25 similar size may co-migrate and appear as a single band on the autoradiogram. If only one of the co-migrating fragments were deleted, such a deletion could be missed. By this approach deletions should be detectable in 50% of affected males, and by careful dosage 30 comparisons in 50% of female carriers. This direct method yields unambigous diagnoses of hemi- and heterozygosity eliminating the need for indirect linkage testing in at least half of all families. The routine

analysis of families at risk for MD using the complete MD cDNA could dramatically decrease the cost and labor involved and greatly increase the accuracy of other diagnostic methods.

If no deletions are found, DNA polymorphism studies as described above are undertaken. These need to include closely linked flanking probes or segments near the 5' and the 3' ends of the gene as well as intragenic markers as argued above.

5

Following is an example of the use of cDNA 10 probes. Family A (Fig. 10) represents a common clinical situation. That is, having only one male affected with DMD (III-1). The carrier status of his mother (II-2) was unknown, since her creatine kinase levels were in 15 the normal range, as were her mother's (I-2). Her male fetus (III-2) was at risk. A DNA polymorphism study was carried out to determine whether the present fetus (III-2) had inherited the same X-21 region as had his affected brother. The individuals shown in Fig. 10 were 20 tested for 21 DNA polymorphisms detected by 18 different probes, some of which are flanking the MD locus and some were subsequently shown to be derived from the gene itself. No deletions were seen with these probes. Haplotype analysis revealed that the affected male 25 (III-1) had received his Xp21 chromosome region from his grandfather (I-1), transmitted without apparent crossover. Therefore, his grandmother (I-2) should not be a carrier, but the carrier status of his mother (II-2) remained unknown. Either she or her affected son 30 (III-1) could represent a new mutation. Based on the DNA-marker results, the risk of the fetus (III-2) being affected was considered higher than the original 20% estimate. To prevent the abortion of a potentially

unaffected fetus the usefulness of the complete gene cDNA probes of the invention was evaluated for detection of the molecular defect in the affected male.

For cDNA studies, the entire gene was surveyed

for the presence of missing or abnormal-sized

restriction fragments in the affected male (III-1), his
mother (II-2) and controls including his unaffected
maternal grandfather (I-1) from whom the Xp21 region was
derived. Genomic DNA samples were cleaved with HindIII

and Bq1II and hybridized with the cDNA probes
illustrated in Fig. 6 and analyzed by known procedures.

DNA extracted directly from CVS material (III-2) was
digested with HindIII and Bq1II and analysed together
with DNAs from family members (I-1, II-2, III-1),

unrelated male controls (C, M1 - M5) and a normal female
control (F).

The mother (II-2) is apparently homozygous for all the fragments deleted in her son (III-1). The deletion detected with two adjacent segments of the cDNA likely represents the molecular basis of the DMD phenotype in the affected male. Therefore, the fetus "at risk" was determined to be unaffected with a 99% or higher probability.

Mouse MD Probes and Nucleic Acid

25 The above human cDNA was used as a probe in Northern blot analysis of mouse tissue mRNA. The human cDNA detected a 14 kb mouse mRNA species, which is present in very low levels in tissue extracts of mouse newborn leg and gravid (15 day) combined uterus and 30 placenta and in higher levels in newborn heart, adult heart, and adult skeletal muscle.

In these studies, RNA was isolated from mouse tissues by homogenization of frozen, ground tissues in

guanidium thiocyanate followed by pelleting through a CsCl cushion as described in Chirgwin, J.M. et al., Biochemistry 18:5294 (1979).

In order to obtain cDNAs representative of the

mouse MD gene, a cDNA library was constructed from adult
mouse heart as described in Gubler, U. et al., Gene

25:263 (1983) as modified in Hoffman, E.P. et al.,
Science 238:347-350, n.12 (1987), and screened with
probes of cDNA corresponding to the human MD gene.

Seven mouse cDNA clones were obtained from 5 X 10⁵
primary (unamplified) λ gt10 recombinants, using
standard techniques.

The low frequency of MD clones obtained in both human and mouse cDNA libraries suggests that the MD messenger RNA (mRNA) is rare, probably representing about 0.01 to 0.002% of the mRNA in muscle. This relative abundance is in agreement with Northern blot analyses and the frequency with which recombinant DMD clones were obtained from four other mouse and human clones were obtained from four other mouse and human clones were constructed in our laboratory.

Fragments of the mouse cDNA were used as probes for Southern blots of mouse genomic DNA using standard procedures as described in Hoffman et al., Science 238:347-350 (1987). In view of the 14 kb size of the complete mRNA, and assuming a constant ratio of cDNA to genomic DNA, the complete mouse MD genomic locus probably encompasses more than 500 kb of genomic DNA. Thus, the hybridization characteristics of the X-linked human MD cDNA described above are also conserved in the mouse.

The DNA sequences of the first about 4.3 kb mouse MD cDNA was determined (Figs. 7A and 7B) and compared to human cDNA sequence. The total nucleotide

sequence of 4.3 kb represents nearly 30% of the entire DMD mRNA.

Interspecies comparison of the sequences from the 5' end of the mouse and human cDNAs (Fig. 7a) show that the first 200 bp are 80% homologous with multiple insertion/deletion differences and translation stop codons in three reading frames. Furthermore, the first translation initiation codon is conserved in both humans and mice.

- Both the DNA and amino acid sequences are well conserved, exhibiting 88% homology with the DNA and 87% with the amino acid sequence. There is a particularly striking conservation in the hydropathicity profile of the mouse and human amino acid sequences, the
- hydropathicity profiles being nearly identical. These profiles were determined according to Kyte et al., J. Mol. Biol., 157:105 (1982). Indeed, by conservative amino acid substitutions based on hydropathicity values, the mouse and human polypeptides become more than 95%
- 20 homologous.

WO 89/06286

Recombinant Dystrophin and Fragments Thereof

The cDNA and predicted amino acid sequence for the 14 kb human or mouse MD gene can be used to produce recombinant dystrophin, which is named after the

25 muscular dystrophy gene to which it corresponds. The amino acid sequence of dystrophin is shown in Fig. 8.

For example, MD cDNA may be inserted into an appropriate vector, introduced into a cell and transcription and translation of the DNA induced, to produce dystrophin. Similarly, smaller fragments may be produced by using only a part of the total cDNA.

Myoblast cells are a suitable expression system. They are cultured and myotube formation induced by standard

20

techniques, as described by Yasin et al., J. Neurol. Sci., 32:347-360 (1977). Other systems include bacteria, such as <u>E. coli</u>, Yeast, and cultured mammalian cells. In such systems, the dystrophin may be secreted or not and may be linked to any of several standard promoters or terminators. Dystrophin may also be produced by <u>in vitro</u> translation of mRNA encoding dystrophin. It may also be produced synthetically.

The above examples are not intended to limit

the present invention but are merely suggestions as to
how dystrophin and fragments thereof can be produced.

The invention encompasses cDNAs, mRNAs, and polypeptide sequences that are modified but substantially correspond to the cDNA and polypeptide sequences defined herein. A nucleic acid or polypeptide sequence substantially corresponds to a given sequence when it is identical except for conservative nucleotide or amino acid substitutions.

Example 4: Dystrophin Fusion Polypeptides
The DNA and predicted amino acid sequences for
4.3 kb (30%) of the mMD gene described above were used
to produce fusion polypeptides and substantially
purified mouse dystrophin.

Two different regions of the mouse heart mMD

25 cDNA were fused to the 3' terminus of the E. coli trpE

gene using the expression vector pATH2 (Dieckmann et
al., J. Biol. Chem 260:1513-1520 (1985)). The two

regions represent the majority of the mouse heart cDNA

sequence described above and shown in Fig. 7B. One

30 construction resulted in the fusion of approximately 30

kD of mMD dystrophin to the 33 kD trpE polypeptide,

while the second fused roughly 60 kD of the mMD

polypeptide. Since the trpE polypeptide is insoluble,

quantitative yields of induced fusion polypeptides were obtained by lysing the cells and precipitating insoluble polypeptides. Novel insoluble fusion polypeptides of the expected size were produced which were not present in lysates of bacteria containing the pATH2 vector alone.

To produce the trpE + 60 kD fusion, the mouse MD cDNA was restricted at the unique SpeI site, blunt-ended with Klenow, and then digested with HindIII in the 3' polylinker. The excised cDNA fragment of 1.4 kb was gel purified and ligated to pATH2 which had been digested with SmaI and HindIII. Recombinants were identified by colony hybridization and verified by subsequent plasmid DNA restriction analysis. The resulting plasmid construction fused the trpE
polypeptide (33 kd) to 410 amino acids (~60 kD) of the mMD polypeptide, and corresponds to position 1.3 kb to 2.7 kb on the equivalent human cDNA map.

end of the mouse cDNA was restricted at its unique

20 non-methylated XbaI site, and at the BamHI site in the

3' polylinker. The excised 700 bp fragment was ligated
to pATH2 and digested with XbaI and BamHI as described
above. This plasmid construction fused the trpE
polypeptide to 208 amino acids (~30 kD) of the mDMD

25 polypeptide, and corresponds to position 3.7 kb to 4.4
kb on the equivalent human cDNA map.

Antibodies can be raised by standard techniques to any of the above described recombinant dystrophin, or 30 fragments thereof. Further naturally ocurring dystrophin may also be used. The following is an example of such a method and is not limiting to this invention.

Antibodies to Dystrophin or Fragments Thereof

Both of the fusion polypeptides described above were purified by preparative SDS-polyacrylamide gel electrophoresis (Laemmli, Nature 227:680-685 (1970)), and used to immunize rabbits and sheep. Rabbits were immunized with electroeluted, 'native' (free from SDS) insoluble antigen, while sheep were immunized with SDS-polyacrylamide gel slices containing denatured antigen.

produced in each rabbit were constantly monitored by enzyme-linked immuno-assays performed on nitrocellulose dot blots of insoluble polypeptide fractions. The best immune responses were obtained using the trpE + 30 kD polypeptide with >95% of the antibodies produced being specifically against the mDMD portion of the fusion peptide, and with titers greater than the sensitivity of the ELISA assay system when using a 1:1000 dilution of crude serum 4 weeks after immunization. The <a href="mailto:trpE + 60 kD antigen took much longer (12 weeks) to evoke a immune response in rabbits, with the resulting sera showing a low specificity for the MD portion of the fusion polypeptide.

Ito ensure that any polypeptide species identified by the antisera was due to recognition by antibodies specific for the MD portion of the fusion polypeptides, rabbit and sheep antibodies directed against the 30 kD antigen, and sheep antibodies directed against the 60 kD antigen were affinity purified. Affinity purification (AP) of the antibodies directed against the MD portion of the fusion polypeptide is facilitated by the insolubility of the partially purified fusion polypeptide. In general, by simply resuspending crude insoluble trpE polypeptide fractions in immune serum, antibodies against the trpE polypeptide

are eliminated. Antibodies specific for the MD polypeptide are then isolated by binding with fusion These antibodies allow detection of as polypeptides. little as 1 ng of dystrophin in 50 μ g muscle tissue.

For example, approximately 3 mg of a partially purified insoluble fraction of trpE polypeptide is precipitated, resuspended in 10 mM Tris (pH 8.0), and then precipitated again. The pellet is resuspended in 1.5 ml of immune serum, incubated on ice for 1 hour, 10 then centrifuged to pellet the trpE-antibody immune

complexes, which are discarded. The supernatant is then mixed with approximately 3 mg of partially purified fusion polypeptide (insoluble fraction) which has been washed as above. After incubation on ice, the

15 mDMD-antibody immune complexes are precipitated by centrifugation. The pellet is then resuspended in 500 ul of 0.2 M glycine (pH 2.3), incubated on ice for 5 minutes to disassociate the immune complexes, and then centrifuged at 4° C to precipitate the insoluble

20 antigen. The supernatant contains purified anti-MD immunoglobulins and is neutralized with 50 ul Tris (pH 9.5), and either stabilized with BSA (fraction V) 5 mg/ml) or dialyzed extensively against phosphate buffered saline (PBS).

25 The invention also contemplates the production of monoclonal antibodies to dystrophin or fragments thereof. Such monoclonal antibodies are secreted by hybridomas produced by standard techniques. Natural Dystrophin

Naturally occuring dystrophin can be identified 30 using the above-described antibodies. For example, total polypeptide samples were isolated from mouse

WO 89/06286 PCT/US88/04504

(fresh) and human (frozen) tissues by direct solubilization of tissues in ten volumes of gel loading buffer (100 mM Tris pH 8.0, 10% SDS, 19 mM EDTA, 50 mM DTT). Alternatively, Triton X-100 insoluble fractions are isolated from human and mouse tissues by homogenizaton in 0.25% Triton X-100 using a Waring blender at full speed, and pelleting insoluble polypeptides. All polypeptide samples (50 ug) are separated by electrophoresis on 3.5% to 12.5% gradient 10 SDS-polyacrylamide gels (Laemmli, Nature 227:680-685 (1970)) using a 3.0% stacking gel, and transferred to nitrocellulose (Towbin, Proc. Nat. Acad. Sci. USA 76:4350-4354 (1979)). Identical nitrocellulose blots of the separated polypeptides are incubated with affinity 15 purified rabbit antibodies directed against the 30 kD antigen, affinity purified sheep antibodies directed against the 60 kD antigen, and affinity purified sheep antibodies directed against the 30 kD antigen each at a 1:1000 dilution. Immune complexes are detected using 20 either ¹²⁵I-polypeptide A, or alkaline phosphatase conjugated donkey anti-sheep IgG second antibody (sheep IgG binds very poorly to polypeptide A). All antibodies detect a large molecular weight, apparently low abundance polypeptide species calculated to be 25 approximately 400 kD in total. The higher resolution of the alkaline phosphatase staining resolves this polypeptide into doublets or triplets, though the slightly smaller bands most likely represent degradation products since there has been no evidence to date for 30 alternatively spliced isoforms of the MD mRNA. kD species is clearly evident in mouse smooth muscle . (stomach), though at a level that is substantially less than that found in cardiac and skeletal muscle.

same apparent polypeptide species is detectable in mouse brain at an extremely low level.

The 400 kD polypeptide species recognized by the antibodies of the invention is generally Triton-5 insoluble, though it appears to be associated more strongly with the myofibrillar matrix fraction in cardiac muscle than in either skeletal muscle or smooth muscle. The 400 kD polypeptide species is present in the skeletal and cardiac muscle of both normal and Tr 10 mice. The detected polypeptide appears the same with a mixture of antisera as it does with each antiserum seperately, indicating that the antibodies raised against different antigens recognized the same polypeptide. Both antibodies fail to detect the 400 kd 15 polypeptide in muscle tissues isolated from mice harboring either allele of the mdx mutation. Immunodiagnosis

The antibodies described above can be used in a variety of immunological applications to test for the 20 presence of dystrophin or fragments thereof in biological samples. For example, the antibodies can be labelled by conventional procedures with ¹²⁵I, ³⁵S, or ³H for use in radioimunoassays, with fluorescein for fluorescent assays, with enzyme for enzyme immunoassays, or with biotin for biotin-avidin linked assays.

These antibodies can be used labelled or unlabelled as desired and can be employed in competitive immunoassays as well as in double antibody or "sandwich" assays. Monoclonal antibodies can also be used in standard techniques to test polypeptides for cross-reactivity with dystrophin.

The antibodies can also be immobilized on an insoluble phase, such as an insoluble resin. The dystrophin level is then detected by measuring the amount of binding to the insoluble phase. Other

insoluble phases include latex particles, which will agglutinate when coated with the novel antibodies and subjected to certain levels of dystrophin in the sample. Yet other insoluble phases include test tubes, vials, titration wells, and the like to which the antibodies according to the invention can be bound and the dystrophin "antigen" thereto detected by double antibody techniques or Protein A dependent techniques. These antibodies are also useful for purification of dystrophin by affinity purification as described above.

In particular, the antibodies according to the invention can be used for the immunodiagnosis of MD.

Futhermore, these novel antibodies can be used to distinguish between Becker's MD (BMD), Duchenne's MD (DMD), Outlier's MD (OMD) and other neuromuscular disorders. For example, muscle biopsies of patients can be prepared and tested according to the invention to determine with a high degree of accuracy whether the patient has DMD, BMD, OMD, or no MD.

To perform such a test, small portions of the

25 biopsies (less than 0.1 g) are crushed into a very thin
sheet of tissue using a pestle and a plastic weigh boat
on a hard surface at -20°C. The tissue fragments are
then weighed at 4°C, and immediately placed in 20
volumes of loading buffer (10% SDS, 50 mM DTT, 10 mM

30 EDTA, 0.1 M Tris pH 8.0, 0.001% bromophenol blue) and
vigorously shaken. Samples are stored at 4°C prior to
use.

Samples are placed in a boiling water bath for two minutes, then centrifuged to remove SDS-insoluble

polypeptides. Supernatants (1-2 ul; 50 ug) are loaded on 20 well, 0.8 mm x 15 cm x 17 cm SDS-polyacrylamide gels (Laemmli, Nature, 227: 680-85 (1970)), using a 3% (0.1% bisacrylamide) stacking gel and 3.5%-12.5% gradient resolving gel. Fractionated polypeptides are transferred to nitrocellulose (Towbin, H. et al., PNAS, 76: 4350-54 (1979)), then incubated with the affinity purified sheep anti-mouse dystrophin described above. Dystrophin/anti-dystrophin immune complexes are detected using affinity purified donkey anti-sheep IgG second antibody conjugated to alkaline phosphatase (Sigma), followed by colorimetric development. All gels are stained with Coomassie Blue after transfer to determine the efficiency of the transfer.

15 Residual myosin in the post-transfer Coomassie blue-stained gels serves as a control for the amount of muscle tissue contained in the biopsy. The accuracy of this residual myosin quantitation of muscle polypeptide is verified by alpha-actin immunostaining after 20 dystrophin staining is completed, or by calsequestrin immunostaining.

Example 7: Muscle Immunodiagnosis

Samples from 103 muscle biopsies were

analyzed. All samples were scored for the abundance and

apparent size of the polypeptide product of the MD gene,
dystrophin, described above.

The abundance and apparent molecular weight of dystrophin contained in 103 muscle biopsies from various neuromuscular disease patients was determined relative to a normal control (examples are shown in Figure 11). Biopsies obtained from patients previously diagnosed as Duchenne exhibited dramatically reduced levels of dystrophin (Fig. 11a, lanes 1, 2, 9-16; Fig. 11b, lanes

3, 8, 9, 12, 14, 16) relative to the normal control

(Fig. 11a, lane 17; Fig. 11b, lanes 1, 5, 17). On the other hand, most patients previously diagnosed as being afflicted with a neuromuscular disorder unrelated to Duchenne or Becker exhibited dystrophin levels indistinguishable from the normal control (Fig. 11b, lanes 6, 7). The Becker biopsy samples exhibited a more variable dystrophin phenotype (Fig. 11a, lanes 3, 5, 6; Fig. 11b, lanes 2, 4, 10, 11). Some Becker biopsies 10 contained dystrophin of apparently smaller molecular weight but of nearly normal abundance (Fig. 11a, lanes 5, 6). One Becker shown in Figure 11b, lane 2, appeared to have dystrophin which was larger in size, but reduced in abundance to approximately 40% of normal. Another 15 Becker exhibited a dystrophin phenotype which was undistinguishable from normal (Fig. 11b, lane 4), while an additional Becker contained no detectable dystrophin in his biopsy (Fig. 11a, lane 3). Three 'Outlier' biopsies (mild Duchenne/severe Becker) are also shown, 20 two of which exhibit dystrophin of low abundance (Fig. 11a, lane 11; Fig. 11b, lane 15), while the third contains no detectable dystrophin (Fig. 11b, lane 13). The size variations of dystrophin were often subtle, such that it was difficult to unambiguously 25 determine size differences if the sample in question was not run immediately adjacent to normal controls (see two Becker biopsies, Fig. 11b, lanes 10, 11). For this reason, all samples containing detectable dystrophin determined by the above analysis were retested adjacent 30 to normal controls. An example of this analysis is

contain dystrophin which was either clearly smaller

(lanes 2, 6) or larger (lane 4) than that of the normal

The three Becker biopsies shown

shown in Figure 12.

controls (lanes 1, 3, 5).

The described experimental process was reported for each of the 103 biopsies included in the study. dystrophin molecular weight (normal vs. abnormal size), and abundance relative to the normal control sample was 5 thus calculated for each patient's biopsy. Five dystrophin 'phenotypes' were established which accounted for all patient dystrophin data, and which seemed to correlate to disease type and severity. dystrophin phenotypes ranged from 'normal' (normal size, 10 60-100% of normal abundance), to 'severe' (less than 3% of normal abundance). The number and percentage of patients under each of four clinical categories (Duchenne, Outliers, Becker, non-DMD/BMD) was then tabulated with regards to the five dystrophin phenotypic 15 classes (Table 2).

Ninety-five percent (38/40) of the patients previously diagnosed as having disorders unrelated to Duchenne or Becker muscular dystrophy exhibited completely normal dystrophin phenotypes. On the other 20 hand, 92% (58/63) of the patients previously diagnosed as having a Duchenne/Becker dystrophy exhibited a clearly abnormal dystrophin phenotype. Within the Duchenne/Becker individuals exhibiting an abnormal dystrophin phenotype there was a clear correlation . 25 between the severity of the clinical phenotype and the 'severity' of the dystrophin phenotype (Fig. 13, roman numerals correspond to categories in Table 2). From this analysis it becomes apparent that the majority of Duchenne patients exhibit undetectable levels of 30 dystrophin, the Outliers (mild Duchenne/severe Beckers; included two affected females) exhibit low levels of . dystrophin of normal size, while most Beckers exhibit nearly normal levels of dystrophin of abnormal size.

- 36 -

There are individual patients whose clinical phenotype clearly does not correlate with the dystrophin phenotype (Table 2, asterisks). These patients comprised 7.7% of the total patients studied (8/103).

5

Table 2

Results of dystrophin analysis in 103 patient biopsies.

Clinical Diagnoses

30	Totals	38	7	18	40
	V. <3% level or undetectable	35 (92.1%)	1 (14.3%)	1 (5.5%)*	
25	IV. Abnormal size 3-40% level	, 1 (2.6%)	1 (14.3%)		
20	III.Normal size, 3-60% level	1 (2.6%)	4 (57.1%)	1 (5.5%)	
	II. Abnormal size	,	1(14.3%)	12 (66.6%)	2 (5.0%)*
15	I. Normal size, 60-100% level	1(2.6%)*		4 (22.2%)*	38 (95.0%)
	Dystrophin Data				
10		<u>Duchenne</u>	<u>Outliers</u>	Becker	Non-DMD/BMD

Therapy

5

The dystrophin or biologically functional fragments thereof can also be used in a method of therapy for MD. A biologically functional polypeptide fragment is capable of causing the same biological effects as the normal full polypeptide. The method of therapy involves administering an amount of the dystrophin or biologically functional fragments thereof effective to control muscle fiber degeneration and 10 connective tissue proliferation within muscle fibers.

The dystrophin and related polypeptides as described above can be administered by routine methods. For example, they can be injected directly into the blood stream or muscle of an animal or human patient to 15 a final concentration of between 1 and 500 mg/ml of serum, most preferably 200 mg/ml. This dose is repeated to maintain this level.

A further method of therapy involves removing cells, such as muscle cells, from an MD patient and 20 incorporating MD DNA into them. The 14 kb cDNA of the invention, or a fragment thereof, is linked to a suitable promoter and inserted into a suitable vector. The vector is then introduced into the cells harvested from the patient and the genetically manipulated cells 25 induced to produce the dystrophin polypeptide, or biologically functional fragments thereof. These cells are then reintroduced into the patient, for example, intravascularly or intramuscularly.

Deposit

30

The complete 14kb cDNA sequence depicted in Figure 5 has been deposited in the form of 6 cDNA probes depicted in Figure 6a. These 6 probes have been deposited with the ATCC under the access numbers 57666

WO 89/06286 PCT/US88/04504

- 38 -

to 57677, inclusive. The deposit was made on July 15, 1987. Each probe is deposited as a purified plasmid and as a plasmid in <u>E. coli</u>. Hence, each probe has two access numbers.

Applicants' assignee, Children's Medical Center Corporation, acknowledges its responsibility to replace these cultures should they die before the end of the term of a patent issued hereon, 5 years after the last request for a culture, or 30 years, whichever is the longer, and its responsibility to notify the depository of the issuance of such a patent, at which time the deposits will be made irrevocably available to the public. Until that time the deposits will be made available to the Commissioner of Patents under the terms of 37 CFR Section 1.14 and 35 Section 112.

Other embodiments are within the following claims.

5

Claims

- 1. An MD probe comprising a substantially purified single-stranded nucleic acid sequence capable of selectively hybridizing to a region of DNA on a human X chromosome between the deletion break point at Xp21.3 and the translocation break point at X;11.
- 2. The MD probe of claim 1 comprising a nucleic acid sequence of at least 10 bases in length from a portion of the sequence of a cDNA probe selected from the group consisting of cDNA probes deposited with the A.T.C.C. under the twelve access numbers 57666 to 57677, inclusive.
- 3. The MD probe of claim 1 wherein said nucleic acid sequence is incapable of selectively

 15 hybridizing to the region of human DNA between said deletion breakpoint at Xp21.3 and the p terminus of said X chromosome, and to the region of DNA between said translocation point at X;11 and the centromere of said X chromosome.
- 20 4. An MD probe comprising a substantially purified single-stranded nucleic acid sequence capable of selectively hybridizing to at least a part of the MD gene.
- 5. The MD probe of claim 4 wherein said nucleic acid sequence is incapable of selectively hybridizing to the regions outside of said MD gene.
 - 6. The MD probe of one of claims 1-5 wherein said probe sequence is capable of hybridizing to at least part of the human gene encoding dystrophin.
- 7. A method for diagnosing MD in a human patient comprising exposing the probe of any one of claims 1 to 5 to a sample taken from said patient and detecting whether said probe hybridizes to nucleic acid in said sample.

8. The method of claim 7 further comprising the steps of,

obtaining a sample of genomic DNA from said patient,

digesting said DNA with one or more restriction endonucleases,

denaturing said DNA,

contacting said denatured, digested DNA with a probe which comprises a nucleic acid sequence at least 10 15 bases in length from the human MD gene,

detecting hybrid DNA complexes, and comparing the molecular weight of said complexes to control complexes formed with similarly denatured and digested genomic DNA taken from a person not suffering from MD,

whereby the presence or absence of deletion mutations in the DMD gene is determined.

- 9. The method of claim 8, wherein said probe comprises one or more probes which in combination span 20 said human MD gene.
 - 10. The method of claim 9 wherein said probes are chosen from those deposited with the A.T.C.C. under the twelve accession numbers 57666 to 57677, inclusive.
 - 11. Substantially pure dystrophin polypeptide.
- 25 12. The polypeptide of claim 11 wherein said dystrophin is human dystrophin.
 - 13. The polypeptide of claim 11 wherein said dystrophin is murine dystrophin.
- 14. A substantially pure polypeptide 30 comprising a sequence of at least six amino acid residues of dystrophin.
 - 15. A substantially pure polypeptide immunologically cross-reactive with dystrophin.

- 16. Substantially purified nucleic acid encoding dystrophin or a polypeptide immunologically cross-reactive with dystrophin.
- 17. A polypeptide synthesized from an engineered nucleic acid sequence corresponding to an MD gene or a segment thereof at least 18 base pairs in length.
- 18. DNA comprising cDNA encoding dystrophin or a sequence of said cDNA of at least 10 base pairs in 10 length.
 - 19. Substantially purified nucleic acid comprising an approximately 14 kb pair sequence of MD gene cDNA.
- 20. The nucleic acid of claim 19 wherein said 15 MD gene is a human MD gene.
 - 21. The nucleic acid of claim 19 wherein said MD gene is a murine <u>Dmd</u> gene.
 - 22. Antibodies produced to a polypeptide of any one of claims 11 to 15, or 17.
- 20 23. The antibodies of claim 22 wherein said antibodies are monoclonal antibodies.
 - 24. A method for immunodiagnosis of MD in a human comprising the steps of,

contacting a sample from said human with an 25 antibody of claim 22 or 23, and

detecting the presence in said sample of complexes of antibody and dystrophin or fragments thereof.

- 25. The method of claim 24 wherein said sample 30 comprises muscle tissue.
 - 26. The method of claim 24 further comprising the steps of,

determining the molecular weight of said dystrophin, or fragments thereof, in said complex,

whereby the lack of dystrophin or the presence of dystrophin of abnormal molecular weight is indicative of MD.

27. The method of claim 24 further comprising 5 the steps of determining the abundance of said dystrophin, or fragments thereof, in said complex.

whereby the presence of an undetectable or abnormally low level of dystrophin is indicative of MD.

- 28. The method of claim 26 further comprising comparing said molecular weight of dystrophin or fragments thereof in said complex from said sample to the molecular weight of a polypeptide cross-reactive with said antibody in at least one BMD, DMD or OMD control and a normal control comprising normal dystrophin,
- whereby said method provides a diagnosis of DMD from the presence in said sample of a polypeptide of molecular weight comparable to said DMD control, a diagnosis of BMD from the presence in said sample of a polypeptide of molecular weight comparable to said BMD control, and a diagnosis of the absence of MD from the presence in said samples of a polypeptide of molecular weight comparable to said normal control.
- $\,$ 29. The method of claim 28 further comprising 25 the steps of

determining the molecular weight of a plurality of fragments reactive with said antibody in said tissue sample to obtain a sample molecular weight pattern,

comparing said sample molecular weight pattern 30 to said BMD, DMD and said normal control patterns,

whereby said method provides a diagnosis of DMD in the presence of a pattern comparable to the DMD control pattern, a diagnosis of BMD in the presence of a

5

pattern comparable to the BMD control pattern, and a diagnosis of the absence of MD in the presence of a pattern comparable to said normal control.

- 30. The method of any one of claims 24 to 29 wherein said sample is obtained from a human adult.
- 31. A method of therapy for MD comprising administering an amount of the polypeptide of any one of claims 11 to 15 or 17 claims effective to control muscle fiber degeneration and connective tissue proliferation within muscle fibers.
 - 32. A method of therapy for MD comprising the steps of,

obtaining cells from an MD patient,

inserting nucleic acid encoding dystrophin or a

15 fragment thereof into a vector,

introducing said vector into said cells, and reintroducing said cells into said patient in an amount effective to produce dystrophin sufficient to control muscle fiber degeneration and connective tissue proliferation within muscle fibers within said patient.

33. A method for purification of dystrophin or a fragment thereof, comprising the steps of;

providing an antibody specific for said dystrophin,

binding said antibody to a solid phase,
contacting said solid phase with a solution
comprising said dystrophin, under conditions wherein
said dystrophin is bound to said solid phase, and

washing said solid phase under conditions
30 wherein said bound dystrophin is released from said solid phase.

34. A method for producing dystrophin, or a fragment thereof, comprising the steps of,

providing nucleic acid encoding said dystrophin or said fragment,

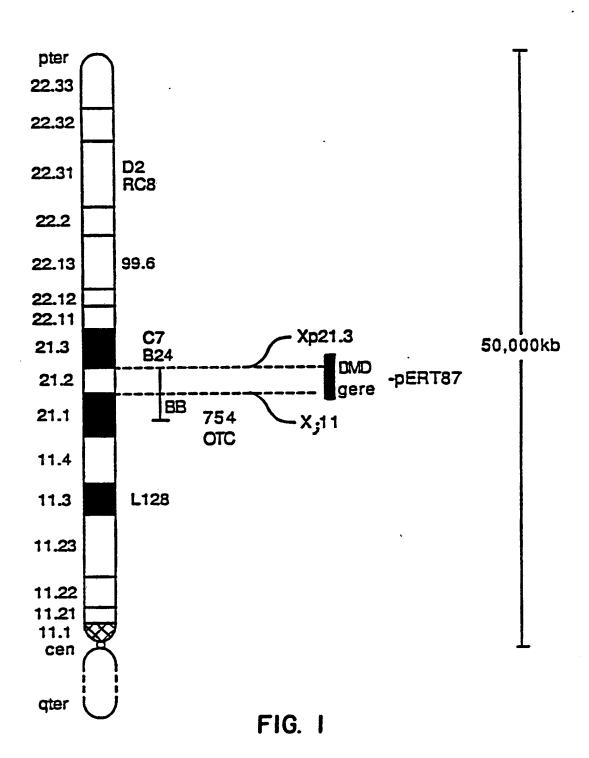
introducing said nucleic acid into a cell, and culturing said cell under conditions suitable to induce expression of said nucleic acid to produce said dystrophin.

- 35. A method of assaying a sample of genomic nucleic acid of a human subject for the presence of a mutated MD gene comprising the steps of,
- denaturing the DNA of said sample and digesting said DNA with one or more restriction endonucleases, contacting said denatured, digested DNA with a probe according to any one of claims 1 to 6, and detecting hybrid DNA complexes as probable indications of the presence of said MD gene in said sample.
- 36. The method of claim 35 wherein said nucleic acid encodes a fusion polypeptide, wherein said dystrophin or said fragment is produced as a polypeptide 20 having a peptide linked to a second polypeptide.
 - 37. The method of claim 36 wherein said second polypeptide comprises a portion of an <u>E. coli</u> polypeptide, and said dystrophin fragment has a molecular weight of between 30 kD and 60 kD.
- 25 38. A method for purification of an antibody, wherein said antibody is specific for a polypeptide, said method comprising the steps of:

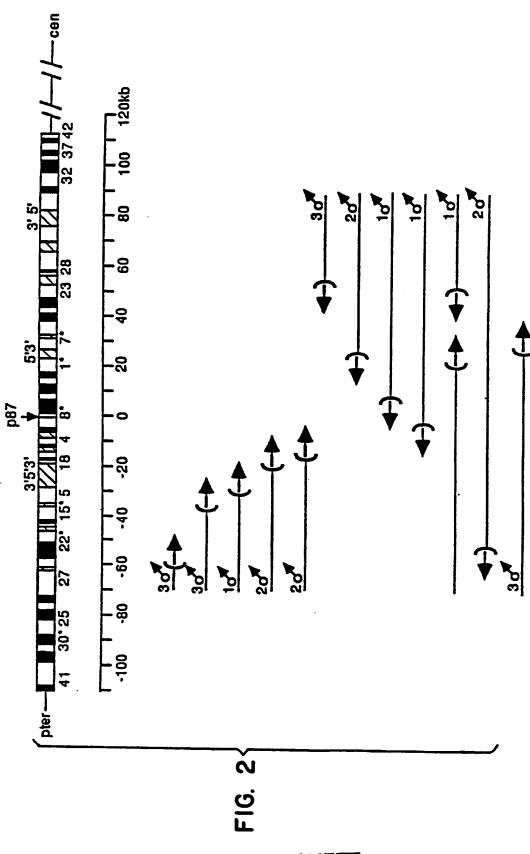
providing said polypeptide in insoluble form, contacting said polypeptide with said antibody 30 to form an immune complex,

5

- 39. The method of claim 38 wherein said polypeptide is fused to an insoluble peptide.
- 40. The method of claim 7 wherein DNA in said sample is fixed to a substrate and exposed to a labelled probe.



SUBSTITUTE SHEET



SUBSTITUTE SHEET

TTCTGTCTTAGTGAAGTCTCAAT-ATTTTTCAATGTGATTAAAATTCATATTATGTCACC

ATAGTTATGAAAAAAGTAAAAATATGTTTTGTATAAGTCTTAT-ATTTGATAATTTAAGC

CAATTAACAAATCCTGACAGCTGGTGTAGAATTTGAGATAAAATTCCCCTGCTGACTCAG

GCAGTGAATAAGAATTTAATGTTCATTAAAATATGTCATAATCTGAGACCCAAAGCAGGT

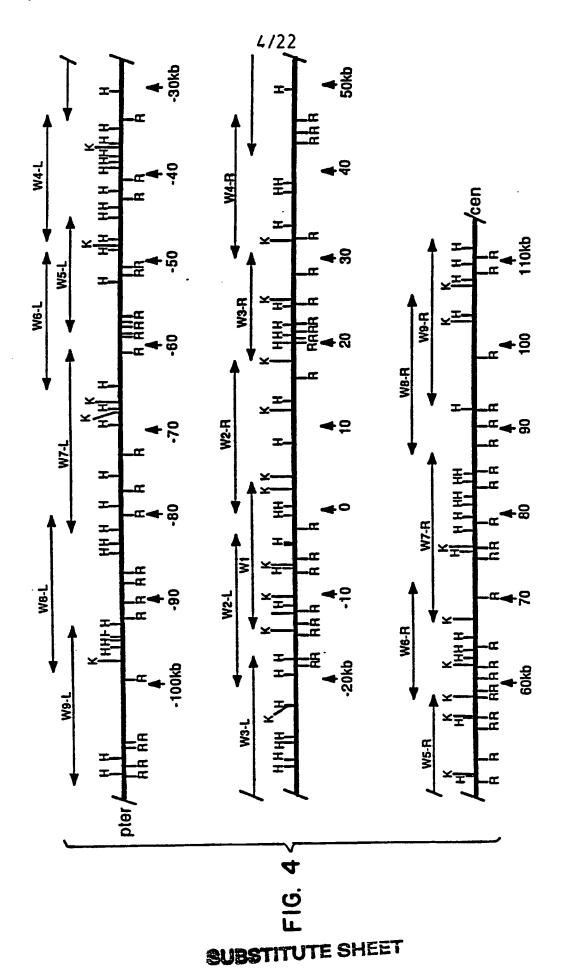
ACCATCTGTTTCCAGTTCCTACTACATGCCTAGGTATGGAATCCATGTTCTAAAATTCTT

TTTAATTTACAACAAGTATAATTTATTTCGAAATAGCATCATCAATAACCAGTGTAATTC

TCGAATATTTTTGTTGGGATTGTTATCTGAGAAATGTGGTATTTCATTTCTTTACAA

FIG. 3

SUBSTITUTE SHEET



ggaagtatgacatcgtagataggaaatcataaactgactottggtttgatttggatataatcctccact GCTTTGGTGGGAAAAGTAGAGGACTGTTATGAAAGAGAAGATGTTCAAAAAGAAAACATTCACAAAATGG GGCAGGTCAAAATGTAATGAAAATATCATGGCTGGATTGCAAGAAACCAAGAGTGAAAAATTCTCCT 3agctgggtgcgacaatcaactcgtaattatccacaggttaatgtaatgaagttgaccaccagctggtct 3GGATTOCCTCACTTTOCCCCTACAGGACTCAGGATCTGGGAGGCAATTACCTTCGGAGAAAAACGAATAG GAAAAACTGAAGTGTTACTTTTTTAAAGCTGCTGAAGTTTGTTGGTTTCTCATTGTTTTAAGGCTACT GTAAATGCACAATTTTCTAAGTTTGGGAAGCAGCATATTGAGAACCTCTTCAGTGACCTACAGGGTGGGGA

FIG. 5 (10F4)

FIG. 5 (20F4)

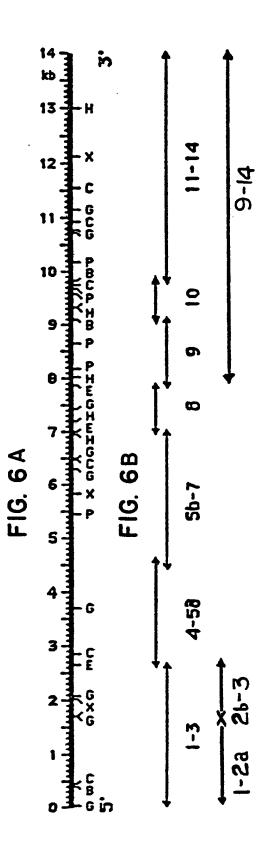
GATGGCCTGGCTTTGAATGCTCTCATCCATAGTCATAGGCCAGACCTATTTGACTGGAATAGTGTGTTT GCCAGCAGTCAGCCACCACGACTGGAACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAA **ACTACTCGATCCTGAAGATGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCA** CTCTTCCAAGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCCAC CTAAAGTGACTAAAGAAGAAGATTTTCAGTTACATCATCAAATGCACTATTCTCAACAGATCACGGTCAG **TCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAAGAGCTATGCCTACACACAGGCT** GCTTATGTCACCACCTCTGACCCTACACGGAGCCCATTTCCTTCACAGCATTTGGAAGCTCCTGAAGACA AGTCATTTGGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGT **ATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTG** GTGAAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGGTA <u>ATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGAAACTGAAGTACA</u> <u>TTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAG</u> AAGAAAGAACAAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACA <u>GTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAG</u> ATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACATCCTTCTCAAATG GCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGAAAAAGAAGATGCAGTGAACAAG **ATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACTGGCCGTTTTAAAAG** CGCATCTAGAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAGAAGATCTTCTTTCAACACT **GAAGAATAAGTCAGTGACCCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTA GTCCAAAAACTTGAAAAGAGTACAGGACAGATTTCACAGGCTGTCACCACCACTCAGCCATCACTAACAC** <u>agagaactetaategaaacagtaactacegtgaccacaaeggaacagatcctegtaaagcatgctcaaga</u> **GGAACTTCCACCACCACCTCCCCAAAAGAAGAGGCAGATTACTGTGGA+TCTGAAATTAGGAAAAGGTTG CATGITGATATAACTGAACTTCACAGCTGGATTACTCGCTCAGAAGCTGTGTTGCAGAGTCCTGAATTTG** CAATCTTTCGGAAGGAAGGCAACTTCTCAGACTTAAAAGAAAAAGTCAATGCCATAGAGCGAGAAAAAGC TGAGAAGTTCAGAAAACTGCAAGATGCCAGCAGATCAGCTCAGGCCCTGGTGGAACAGATGGTGAATGAG **GGTGTTAATGCAGATAGCATCAAACAAGCCTCAGAACAACTGAACAGCCGGTGGATCGAATTCTGCCAGT** TGCTAAGTGAGAGACTTAACTGGCTGGAGTATCAGAACAACATCATCGCTTTCTATAATCAGCTACAACA **GCAATTAAAAGTCAGTTAAAAATTTGTAAGGATGAAGTCAACCGGCTATCAGGTCTTCAACCTCAAATTG AACGATTAAAAATTCAAAGCATAGCCCTGAAAGAGAAAGGACAAGGACCCGATGTTCCTGGATGCAGACTT** TGTGGCCTTTAGAAATGATTTTAAGGAAGTCTTTTCTGATGTGCAGGCCAGAGAGAAAGAGCTACAGACA ATTITTGACACTITGCCACCAATGCGCTATCAGgAGACCATGAGTGCCATCAGGACATGGGTCCAGCAGT CAGAAACCAAACTCTCCATACCTCAACTTAGTCTCACCGACTATGAAATCATGGAgcAGAGActcgGGGA <u>ATTGCAGGC+TTACAAAGTTCTCTGCAAGAGCAACAAAGTGGCCTATACTATCTCAGCACCACTGTGAAA</u> GCTGGAAGAAGCTCTCCTCCCAGCTGGTTGAGCATTGTCAAAAGCTAGAGGAGCAAATGAATAAACTCCG TGGCCtGCCCTTGGGGATTCAGAAATTCTAAAAAAGCAGCTGaaacagtgcagacttttagtcagtgata ttcagacaattcAGCCCAGTCTAAACAGTGTCAATGAAGGTGGGCAGAAGATaaagaatgaagcagagcc agagtttgcttcgagacttgagacagaactcaaagaacttaacactcagtgggatcacatgtgccaacag gtctatgccagaaaggaggccttgaagggaggtttggagaaaactGTAAGCCTCCAGAAAGATCTatcag tgaattacagaaagcagttgaagagatgaagagagctaaagaagaggcccAACAAAAAGAAGCGAAAGTG **AAACTCCTTACTGAGTCTGTAAATAGTGTCATAGCTCAAGCTCCACCTGTAGCACAAGAGGCCTTAAAAA AGGAACTTGAAACTCTAACCACCAACTACCAGTGGCTCTGCACTAGGCTGAATGGGAAATGCAAGACTTT** GGAAGAAGTTTGGGCATGTTGGCATGAGTTATTGTCATACTTGGAGAAAGCAAACAAGTGGCTAAATGAA **GTAGAATTTAAACTTAAAACCACTGAAAACATTCCTGGCGGAGCTGAGGAAATCTCTGAGGTGCTAGATT** CACTICAAAATTICATCCCACATTCAGAGGATAACCCAAATCAGATTCGCATATTGGCACAGACCCTAAC AGATGGCGGAGTCATGGATGAGCTAATCAATGAGGAACTTGAGACATTTAATTCTCGTTGGAGGGAACTA CATGAAGAGGCTGTAAGGAGGCAAAAGTTGCTTGAACAGAGCATCCAGTCTGCCCAGGAGACTGAAAAAT CCTTACACTTAATCCAGGAGTCCCTCACATTCATTGACAAGCAGTTGGCAGCTTATATTGCAGACAAGGT **GGACGCAGCTCAAATGCCTGAGGAAGCCCAGAAAATCCAATCTGATTTGACAAGTCATGAGATCAGTTTA GAAGAAATGAAGAAACATAATCAGGGGAAGGAGGCTGCCCAAAGAGTCCTGTCTCAGATTGATGTTGCAC** <u>AGAAAAATTACAAGATGTCTCCATGAAGTTTCGATTATTCCAGAAACCAGCCAATTTTGAGCTGCGTCT</u> <u>ACAAGAAAGTAAGATGATTTTAGATGAAGTGAAGATGCACTTGCCTGCATTGGAAACAAAGAGTGTGGAA</u> CAGGAAGTAGTACAGTCACAGCTAAATCATTGTGTGAACTTGTATAAAAGTCTGAGTGAAGTGAAGTCTG AAGTGGAAATGGTGATAAAGAGTGGACGTCAGATTGTACAGAAAAAGCAGACGGAAAATCCCAAAGAACT

FIG. 5 (30F4)

CAGCTACAGATATGGAATTGACAAAGAGATCAGCAGTTGAAGGAATGCCTAGTAATTTGGATTCTGAAGT <u>TGCCTGGGGAAAGGCTACTCAAAAAGAGATTGAGAAACAGAGGTGCACCTGAAGAGTATCACAGAGGTGTA</u> GGAGAGGCCTTGAAAACAGTTTTGGGCAAGAAGGAGACGTTGGTGGAAGATAAACTCAGTCTTCTGAATA GTAACTGGATAGCTGTCACCTCCCGAGCAGAAGAGTGGTTAAATCTTTTGTTGGAATACCAGAAACACAT **GGAAACTTTTGACCAGAATGTGGACCACATCACAAAGTGGATCATTCAGGCTGACACACTTTTGGATGAA** TCAGAGAAAAAGAAACCCCAGCAAAAAGAAGACGTGCTTAAGCGTTTAAAGGCAGAACTGAATGACATAC GCCCAAAGGTGGACTCTACACGTGACGAAGCAGCAAACTTGATGGCAAACCGCGGTGACCACTGCAGGAA **ATTAGTAGAGCCCCAAATCTCAGAGCTCAACCATCGATTTGCAGCCATTTCACACAGAATTAAGACTGGA AAGGCCTCCATTCCTTTGAAGAATTGGAGCAGTTTAACTCAAAAAATTGCTTGAACCACGG** <u>agcotgaattoagcaggggggaatotgaaagaggaagacttaataaagatatgaagagagacatga</u> <u>GGGTACTGTAAAAGAATTGTTGCAAAGAGGAGACAACTTACAACAAAGAATCACGAGATGAGAGAAAGAGA</u> <u>GAGGAAATAAAGATAAAACAGCAGCTGTTACAGACAAAACATAATGCTCTCAAGGATTTGAGGTCTCAAA</u> **ATTGATGGGGAATTGCAGAAGAAGAAGGAGGTGAATGCAGTGAGGGAAGCTGAGGGCTTGTCTG** <u>AGGATGGGGCCGCAATGGCAGTGGAGCCAACTCAGATCCAGCTGAGGAAGCGCTGGCGGGAAATTGAGAG</u> CAAATTTGCTCAGTTTCGAAGACTCAACTTTGCACAAATTCACACTGTCCGTGAAGAAACGATGATGGTG **ATGACTGAAGACATGCCTTTGGAAATTTCTTATGTGCCTTCTACTTATTTGACTGAAATCACTCATGTCT** CACAAGCCCTATTAGAAGTGGAACAACTTCTCAATGCTCCTGACCTCTGTGCTAAGGACTTTGAAGATCT CTTTAAGGAAGAGGAGTCTCTGAAGAATATAAAAGATAGTCTAGAAGAAAGCTCAGGTCGGGATTGACATT <u>ATTCATAGCAAGAAGACAGCAGCATTGCAAAGTGCAACGCCTGTGGAAAGGGTGAAGCTACAGGAAGCTC</u> **TCTCCCAGCTTGATTTCCAATGGGAAAAAGTTAACAAAATGTAGAAGGACCGACAAGGGCGATTTGACAG ATCTGTTGAGAAATGGCGGCGTTTTCATTATGATATAAAGATATTTAATCAGTGGCTAACAGAAGCTGAA** CAGTTTCTCAGAAAGACACAAATTCCTGAGAATTGGGAACATGCTAAATACAAATGGTATCTTAAGGAAC TCCAGGATGGCATTGGGCAGCGGCAAACTGTTGTCAGAACATTGAATGCAACTGGGGAAGAAATAATTCA GCAATCCTCAAAAAGAGATGCCAGTATTCTACAGGAAAAATTGGGAAGCCTGAATCTGCGGTGGCAGGAG GTCTGCAAACAGOTGTCAGACAGAAAAAAGAGGCTAGAAGAACAAAAGAATATCTTGTCAGAATTTCAAA **GAGATTTAAATGAATTTGTTTTATGGTTGGAGGAAGCAGATAACATTGCTAGTATCCCACTTGAACCTGG** <u>aaaagagcagcaactaaaagaaaagcttgagcaagtcaagttactggtggaagagttgcccctgcgccag</u> <u>GGAATTOTCAAACAATTAAATGAAACTGGAGGACCCGTGCTTGTAAGTGCTCCCATAAGCCCAGAAGAGC</u> <u>AAGATAAAOTTGAAAATAAGCTCAAGCAGACAAATCTCCAGTGGATAAAGGTTTCCAGAGCTTTACCTGA</u> <u>GAAAGAAGGAGAAATTGAAGCTCAAATAAAAGACCTTGGGCAGCTTGAAAAAAAGCTTGAAGACCTTGAA</u> <u>ACCAAGAAGGACCATTTGACGTTCAGGAAACTGAAATAGCAGTTCAAGCTAAACAACCGGATGTGGAAGA</u> GATTTTGTCTAAAGGGCAGCATTTGTACAAGGAAAAACCAGCCACTCAGCCAGTGAAGAGGAAGTTAGAA GATCTGAGCTCTGAGTGGAAGGCGGTAAACCGTTTACTTCAAGAGCTGAGGCCAAAGCAGCCTGACCTAG CTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCAGACTGTTACTCTGGTGACACAACCTGTGGTTAC **TAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTCCTTGATGTTGGAGGTACCTGCTGTGGCAGAT** TTCAACCGGGCTTGGACAGAACTTACCGACTGGCTTTCTCTGCTTGATCAAGTTATAAAATCACAGAGGG TGATGGTGGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCAGAAGGCAACAATGCAGGATTTGGA <u>ACAGAGGCGTCCCCAGTTGGAAGAACTCATTACCGCTGCCCAAAATTTGAAAAACAAGACCAGCAATCAA</u> GAGGCTAGAACAATCATTACGGATCGAATTGAAAGAATTCAGAATCAGTGGGATGAAGTACAAGAACACC **TTCAGAACCGGAGGCAACAGTTGAATGAAATGTTAAAGGATTCAACACAATGGCTGGAAGCTAAGGAAGA AGCTGAGGAGGTCTTAGGACAGGCCAGAGCCAAGCTTGAGTCATGGAAGGAGGGTCCCTATACAGTAGAT GCAATCCAAAAGAAAATCACAGAAACCAAGCAGTTGGCCAAAGACCTCCGCCAGTGGCAGACAAATGTAG** atgtgggaaatgacttggccctgaaacttctccgggattattctgcagatgataccagaaaagtccacat GAAGAAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAG CTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGT **AAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAAC** CTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGAC **GTTTGGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGA** <u>agccagttctgaccagtegaagcgtctgcacctttctctgcaggaacttctggtgtggctacagctgaaa</u> GATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATCTAC ATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAAT <u>ATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAG</u> GAGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAT TGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAAGA GGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGC

FIG. 5 (40F4)

GATCTCCTCATTGACTCTCCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTC TGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACC GTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGA AGGGTCCCTGGGAGAGAGCCATCTCGCGAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAAC TTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCA **GCTTATAGGAOTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAG** CTGCATGTGATGCCTTGGACCAGCACAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTAT TAATTGTTTGACCACTATTTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGC GTGGATATGTGTCTGAACTGGCTGCATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGT CTITIAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCA <u>AGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTCCTTCTGCATGATTCTATCCAAATT</u> CCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCC <u>AATTTGCTAATAATAAGCCAGAGATCGAAGCGCCCCTCTTCCTAGACTGGATGAGACTGGAACCCCAGTC</u> CATGGTGTGGCTGCCCGTCCTGCACAGAGTGGCTGCTGCAGAAACTGCCAAGCATCAGGCCAAATGTAAC **ATCTGCAAAGAGTGTCCAATCATTGGATTCAGGTACAGGAGTCTAAAGCACTTTAATTATGACATCTGCC** AAAGCTGCAAAAAAACTGGTCGAGTTGCAAAAGGCCATAAAATGCACTATCCCATGGTGGAATATTGCAC TCCGACTACATCAGGAGAAGATGTTCGAGACTTTGCCAAGGTACTAAAAAACAAATTTCGAACCAAAAGG TATTTTGCGAAGCATCCCCGAATGGGCTACCTGCCAGTGCAGACTGTCTTAGAGGGGGACAACATGGAAA TGATACTCATTCACGCATTGAACATTATGCTAGCAGGCTAGCAGAAAATGGAAAACAGCAAATGGATCTTAT CTANATGATAGCATCTCTCTAATGAGAGCATAGATGATGAACATTTGTTAATCCAGCATTACTGCCAAA GTTTGAACCAGGACTCCCCCCTGAGCCAGCCTCGTAGTCCTGCCCAGATCTTGATTTCCTTAGAGAGTGA **GGAAAGAGGGGAGCTAGAGAATCCTAGCAGATCTTGAGGAAGAAAACAGGAATCTGCAAGCAGAATAT** GACCETCTAAAGCAGCAGCACCATAAAGGCCTGTCCCCACTGCCGTCCCCTCCTGAAATGATGCCCA CCTCTCCCCAGAGTCCCCGGGATGCTGAGCTCATTGCTGAGGCCAAGCTACTGCGTCAACACAAAGGCCG CCTGGAAGCCAGGATGCAAATCCTGGAAGACCACAATAAACAGCTGGAGTCACAGTTACACAGGCTAAGG CAGCTGCTGGAGCAACCCCAGGCAGAGGCCAAAGTGAATGGCACAACGGTGTCCTCCTCCTTCTACCTCTC TACAGAGGTCCGACAGCAGTCAGCCTATGCTGCTCCGAGTGGTTGGCAGTCAAACTTCGGACTCOATGGG TGAGGAAGATOTTCTGAGTCCTCCCCAGGACACAAGCACAGGGTTAGAGGAGGTGATGGAGCAACTCAAC **AACTCCTTCCCTAGTTCAAGAGGAAGAATACCCCTGGAAAGCCAATGAGAGAGGACACAATGTAGGAAG** TCTTTCCACATGGCAGATGATTTGGGCAGAGCGATGGAGTCCTTAGTATCAGTOATGACAGATGAAGAA **GGAGCAGAATAAATGTTTTACAACTCCTGATTCCCGCATGGTTTTTATAATATTCATACAACAAAGAGGA** TTAGACAGTAAGAGTTTACAAGAAATAAATCTATATTTTTGTGAAGGGTAGTGGTATTATACTGTAGATT TCAGTAGTTTCTAAGTCTGTTATTGTTTTGTTAACAATGGCAGGTTTTACACGTCTATGCAATTGTACAA **AAAAGTTATAAGAAAACTACATGTAAAATCTTGATAGCTAAATAACTTGCCATTTCTTTATATGGAACGC** <u>ATTITGGGTTGTTTAAAAATTTATAACAGTTATAAAGAAAGATTGTAAACTAAAGTGTGCTTTATAAAAA</u> **ACTITGAGGCAGCGCATTGTTTTGCATCCTTTTGGCGTGATATCCATATGAAATTCATGGCTTTTTCTTT** TTTTGCATATTAAAGATAAGACTTCCTCTACCACCACACCAAATGACTACTACACACTGCTCATTTGAGA **AGATCCTAATAGAAAGAAATTACTTCTAGTCAGTCATCCAGGCTTACCTGCTTGGTCTAGAATGGATTTT** TCCCGGAGCCGGAAGCCAGGAGGAAACTACACGACACTAAAACATTGTCTACAGCTCCAGATGTTTCTCA **ATTCAGCTTCCTGCTAGCAATGCCACGATTTAGATTTAATGATGCTTCAGTGGANNTCAATCAGAAG** GTATTCTGACCTTGTGAACATCAGAAGGTATTTTTTAACTCCCCAAGCAGTAGCAGGACGATGATAGGGC TGGAGGGCTATGGATTCCCAGCCCATCCCTGTGAAGGAGTAGGCCACTCTTTAAGTGAAGGATTGGATGA TTGTTCATAATACATAAAGTTCTCTGTAATTACAACTAAATTATTATGCCCKCTTCTCACAGTCAAAAGG AACTGGGTGGTTTGGTTTTTGNNNNTTTTTTAGATTTTTTGTCCCATGTGGGATGAGTTTTTAAATGCCA CAAGACATAATTTAAAATAAATAAACTTTGGGAAAAGGTGTAAGACAGTAGCCCCATCACATTTGTGATA CTGACAGGTATCAACCCAGAAGCCCATGAACTGTGTTTCCATCCTTTGCATTTCTCTGCGAGTAGTTCCA CACAGGTTIGTAAGTAAGTAAGAAAGAAGGCAAATTGATTCAAATGTTACAAAAAAACCCTTCTTGGTGG **ATTAGACAGGTTAAATATATAAAGAAAGAAACAAAATTGCTC**



SUBSTITUTE SHEET

										* *	*			-						
HUMAN MOUSE									GG(JAT'	rcc(CTC	ACT?	rtc(rgc(CTA TTA	CAG(GAC'	TCAG TCAG	CTC
					t # :		***	* *					*	**				*		*
•	TGG	• •3 <i>770</i>	ירא	יים חית מ	ים כי				ΔΔΔ	ACG	ል <mark>አ</mark> ጥ	AGG.	AAA	AAC!	rga.	AGT	GTT	ACT"	TTTT	TT-
	TTG	AAGO	3CA	ATA	GC:	rtt	ATA	GAA	AAA	ACG	AAT	AGG	AAG?	ACT	rga.	AGT	GCT	ATT'	TTTI	TTT
	***	***	**	**	*					*		*		*	*	* *	330	a com	★ አ <i>ር</i> መር	CAG
				P	AA(GCT(CT(GAA	GTT.	rgt:	rgg:	Chale Lalale	CTC	ያ.ፐ.ፕ.(31.1. 31.1.	CCIL T.T.T.	AAG DDG	CCT	CCTO	GAG GAG
	TTTT	rtt.	ľľG	TCA	AG	GCT(3CT(JAA	<i>5</i> 1"1".	LAT.	TGG	U11'		AIC(JIN				00-	
			**	*	*	*				*	*		**	*						***
	CAA	[AA]	AGT	TTC	AA	G-A	ACT	TTT	ACC	AGG'	TTT'	\mathbf{rrr}	TTA:	rcg(CTG CTG	CCT	TGA TGA	TAT.	ACA(ATA(.h.h.h.
	CAA!	[AA]	AAC	TGG	GA (GAA	ACT	TTT	ACC	AAG	ATT.	LTT	д.	ICC	CIG	CCI	IOA			CTTT
	*	1	k															- 45	mami	
201	-TT	rtc?	AA A	ATO	CT'	TTG	GTG	GGA	AGA	AGT	AGA	GGA	CTG:	TTA:	IGA	AAG	AGA E	AGA D	TGT.	rcaa Q
					L	W	W	E	E E	V V	E E	ע	C	Y	_	R R	E	D	v	ğ
212	ւերդեր	ኮጥሮሪ	מ מי־	M ATO	L CT	W TTG	vi GTG	GGA.	AGA.	AGT:	AGA	GGA	CTG'	TTA'		AAG	AGA	AGA	TGT:	CAA
212	LII.	116	J. 14 :																	
																	*		* :	k *
	N N G	מממ	מים	יחיתי	יאר:	ΔΔΔ	ልጥG	GGT.	AAA	TGC	ACA	ATT	TTC	TAA	GTT	TGG	GAA	GCA	GCA!	TTAT
				F	T	K	W	V	N	A	Q	F	S	K	r	G	K	ă	п	<u>.</u>
	72 1	K :	r	F	T	K	W	I	N	A	Q	F	S	K	F	G	K מממי	Q	H	I מיזער
	AAG	AAA	ACA	TTC	CAC	AAA	ATG	GAT:	AAA	TGC	ACA	ATT	TTC	TAA	GII	166	mnn	.GCA		CATA
	*							*			* *						*	*		*
	GAG	AAC	CTC	TT:	CAG	TGA	CCT	ACA	GGA	TGG	GAG	GCG	CCT	CCT	AGA	CÇI	'CÇT	CGA	AGG(G	CCTG L
ſ	E			F	S	_	L		D		R K		L	L	D D	L L	L	E E	G	_
i	D	N .	L CTC	F TOTA	S CAG	D TGD	L CCT	Q GCA	D GGA	G TGG	AAA	ACG	CCT	CCT				GGA	AGG	CCTT
	GAC	anc.		. 1 1 \	مم	1011		٠												
										*	3.00	* ''''	יאאכ	አርጥ	מ ייניי	ጥርር	ירריו	GAD	CAA	TGTC
					ACT L	GCC P	aaa K	AGA E	aaa K	AGG G	S	T	R	V	H	A	L	N	N	V
	m .	~ 7	~	W	T	TD.	K	F	ĸ	G	S	T	R	V		A	L	N	N	_ V
	ĀCA	GGG	CAA	\AA	ACT	GCC	AAA	AGA	AAA	.GGG	ATC	TAC	:AAG	AGT	TCP	\TG(CCCI	'GAP	CAA	TGTC
							*	*		*								*		*
	AAC	AAG	GCI	CT	GCG	GGT	TTT	GCA	GAA	CAA	TAA	TGI	TGA	TTT	'AG'I	rga,	LAT!	TGG	BAAG	TACT
	N		Α	L	R	V	L	\sim	I NT	! NT	N	v	n	1.	v	N	1	G	3	1
	N	K .	A	L_	R	V	L	Q	LK.	N	N	V	D Total	anma L	V ים מי	N מבסים	ב דביים	G בי		
	AAC	AAG	GC	ACT(GCG	GGI	CIT	ACA	GAA		(TPM	7101	LIGA	71 1 1	יטה	. 0				CACT
		*	•	b						*										
500	GAC	ATC	GT!	\GA'		AAA	TCP	TAP	ACT	'GAC	TCI	TG	TŢT	'GAI	TT!	3GAI	ATA!	CAAT I	rcct t.	CCAC
	_	_	V	D	G	N	H	K K			L	G	L	I			I		_	H
512	D GNC	I מיזימי	V GT(D CA	G ጥርር	N aaa:	ሊጥር ፤ አ	A LTAD	NVC.	CAC	TCT	G TG(u TTT	[GA]	TT(GGA.	ATA:			CCAC
312	GAC	WTW	.G1(JUM	100	<u></u> .														

FIG.7A-1

QUBSTITUTE SHEET

							*		*							*					
	TG	GCA	GGT	'CAF	AAA	TGT	'AA'I	'GAA	AAA	TAT	CAT	GGC	TGG	ATT	'GCA	ACA	AAC	CAA	CAG	TGAA	L
	W	0	V	K	N	V	M	K	N	I	M	Α	G	L	Q	Q	T	N	S	Ē	
				ĸ		Ÿ		ĸ	T	T	M	A	G	L	Q	Õ	T	N	S	E	
	mc	~~z		א מיטו	777	TO TO	ייי דע בעי	ממכוי	770	ጥልጥ የ	יייע	יבפר	ידינים	יויים ע	יכרא יכרא					TGAA	L
	16	GLA	GGT	CAA	TATA:	77.01	GAI	. GAA	mnc			000	.100	42± ÷	· ·			U			•
				_				•		*		_							*	±	
				*											maa		m	m 2 2	 mam	 3 3 M/C	
	AA	GAT	TCI	CC1				'CCG	ACA	ATC	AAC	TCG	TAA	TTA	TCC	ACA	GG.T.	TAA	TGT	AATC	•
	K	Ι	L	L	S	W	V	R	Q	S	T	R	N	Y	₽	Q	V	N			
	K	I	L	L	S	W	V	R	Q	S	T	R	N	Y	P		V				
•	AA	GAT	TCT	TCI	'GAG	CTG	GGT	TCG	ACA	GTC	AAC	ACG	TAA	TTA	TCC	ACA	GGT	TAA	CGT	CATC	,
				*	*			*	*	**						*				*	
	22	ىلىش	C N C	יראכ	ממי	יידים	ረጥር	ጥርል	TGG	ССТ	CCC	արդու	GAA	TGC	тст	CAT	CCA	TAG	TCA	TAGG	;
					•	W	O T C	יבט. ח	G	T	A	T.	N	A			H	S	Н	R	
	N	_	_	T			2	ע	G	Ţ	M.	Ţ	74	Â	L	Ī	H	S	H	R	
	N	F	T'	<u>s</u>	S	W	8	ע	-6	_L	A	———	<i>V</i>	maa							
	AA	CTT	CAC	CTC	TAG	CTG	GTC	:CGA	CGG	GTT	GGC	TTT	GAA	TGC	TCI	TAT	CCA	TAG	ICA	CAGG	,
1	k		×		*					*	*		*			*	*				
	CC	AGA	CCT	ATI	TGA	CTG	GAA	TAG	TGT	GGT	TTG	CCA	<u>GCA</u>	GTC	AGC	CAC	ACA	ACG	ACT	GGAA	L
			L			W		S	V	V	C	Q	Q	S	Α	T	Q	R	L	E	
	P				Ď	W	N		V				Ĥ			T	Q	R	L	E	
	CC	CGD		ट्र ा ग्ग		ייים. סיניים	227												ACT	GGAA	Ĺ
		CGA	LUI	GII	, TON	.110	0211	1170	101	-											
			_																		
			* 									~~ m		~~ ~	3 00	13 OM	<i>~~</i>	maa	መሮአ	አራክጥ	•
300					CAT	CGC	CAG	ATA	TCA	ATT	AGG	CAT	AGA	GAA	ACT	ACT	CGA	ICC	IGN	AGAT	•
	H	A	F	N	I	A	R	Y	Q	L	G	1	F.	K	L	L	ע	P	E	D	
	H	A	F	N																	
312	CA	TGC	CTT	CAA	CA																
-												-									
	CT	тса	ጥልሮ	CAC	מידי	TCC	AGA	TAA	GAA	GTC	CAT	CTT	'AAT	GTA	CAT	CAC	ATC	ACT	CTT	CCAA	L
		D.			Y			ĸ	ĸ	S	T	Τ.	M	Y	I		S				
	٧	ט	-	_	-	-		••	••	~	•	-	••	-	_	_	_	_		~	
		mmm		<i></i>	303	3 CM	CR C	~ N (T)	መረን	700	יית איי	רכז	CCN	አርጥ	CCI	א מ מ	للسك	ccc	מממ	GCCA	
	GT	111	GCC	TUA	MCA	MGT	GAG	- T	TOW	AGC	T	~~	ייסטי	77	- E	M	L	סטט	סיני	P	١
	V	L	P	Q	Q	V	S	Ţ	E	A	1	Q	c	٧	£	ĽΣ	1.	=	K	F	
	CC	TAA	AGT	'GAC	TAA	aga	aga	ACA	TTT	TCA	GTT	'ACA	TCA	TCA	AAT	'GCA	CTA	TTC	TCA	ACAG	į
	P	K	V	T	K	E	E	H	F	Q	L	H	H	Q	M	H	Y	S	Q	Q	
	ΔT	ראר	ርርጥ	CAG	PCT	AGC	ACA	GGG	ATA	TGA	GAG	AAC	TTC	TTC	CCC	TAA	.GCC	TCG	ATT	CAAG	į
	I	~ TP	77	~~	T.	A	<u> </u>	G	Y	F	R	T	S	S	P	ĸ	P	R	F	K	
	т_	7	٧	3	73	-	×	•	•	_		•	_	_	-		-		_		
1		-			03.0		000	maa	ת יחח	m/m	~ * ~	~~~		መረ እ	CCC	ም ክ ፖ	a cc	CAC	ccc	ATTI	1
1100					CAC	ACA Q	النائنانا.	TGC	TIM	TGT	LAC	.CAC	.616	TOM		TVC	ハしら	ono		. E.	,
	S	Y	A	Y	T	Q	A	A	¥	V	T	T	5	ע	צ	T	K	3	P	£	
								_												as	
	CC	TTC	ACA	.GCA	TTT	'GGA	AGC	TCC	TGA	AGA	CAA	GTC	ATT	TGG	CAG	TTC	ATT	GAT	GGA	GAGT	
	P	S	O	H	L	E	A	P	Ē	D	K	S	F	G	S	S	L	M	E	S	
	_	_	~		_	_															

FIG.7A-2

EVNLDRYQTALEEVLSWLLS GCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGAC A E D T L Q A Q G E I S N D V E V V K D CAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGGT Q F H T H E G Y M M D L T A H Q G R V G 1400 AATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGAA NILQLGSKLIGTGKLSEDEE ACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGCT TEVQEQMNLLNSRWECLRVA AGCATGGAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACTG S M E K Q S N L H R V L M D L Q N Q K L AAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAAAATGGAGGAAGAG KELNDWLTKTEERTRKMEEE CCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCAA P L G P D L E D L K R Q VAQ Q H K V L Q 1699

FIG.7A-3

Mouse 1 AGAGTTATGCCTTCACACAGGCTGCTTATGTTGCCACCTCTGATTCCACACAGAGCCCCT S Y A F T Q A A Y V A T S D S T Q S P Y **ATCCTTCACAGCATTTGCAAGCTCCCAGAGACAAGTCACTTGACAGTTCATTGATGGAGA** PSQHLEAPRDKSLDSSLME CGGAAGTAAATCTGGATAGTTACCAAACTGCTTTAGAAGAAGTACTTTCATGGCTTCTTT EVNLDSYQTALEEVLSWLLS CTGCCGAGGATACATTGCGAGCACAAGGAGAGATTTCAAATGATGTTGAAGAAGTGAAAG A E D T L R A Q G E I S N D V E E V K E **AACAGTTTCATGCTCATGAGGGATTCATGATGGATCTGACATCTCATCAAGGACTTGTTG** Q F H A H E G F M M D L T S H Q G L V G **GTAATGTTCTACAGTTAGGAAGTCAACTAGTTGGAAAAGGGAAATTATCAGAAGATGAAG** N V L Q L G S Q L V G K G K L S E D E E **AAGCTGAAGTGCAAGAACAAATGAATCTCCTAAATTCAAGATGGGAATGTCTCAGGGTAG** A E V Q E Q M N L L N S R W E C L R V A CTAGCATGGAAAAACAAAGCAAATTACACAAAGTTCTAATGGATCTCCAGAATCAGAAAT S M E K Q S K L H K V L M D L Q N Q K L TAAAAGAACTAGATGACTGGTTAACAAAAACTGAAGAGAGAACTAAGAAAATGGAGGAAG KELDDWLTKTEERTKKMEEE Mouse **NAGCCCTTTGGACCTGATCTTGAAGATCTAAAATGCCAAGTACAACAACATAAGGTGCTT** PFGPDLEDLKCQVQQHKVL **▼1** ACAACAACATAAGGTGCTT Human Q E D L E Q E Q V R V N S L T H M V V V Ō E D L E Ō E Ō V R V N S L T H M V V * GTTGATGAATCCAGCGGTGATCATGCAACAGCTGCTTTGGAAGAACAACTTAAGGTACTG V D E S S G D H A T A A L E E Q L K V L V D E S S G D H A T A A L E E Q L K V L **GTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTG** GGAGATCGATGGGCAAATATCTGCAGATGGACTGAAGACCGCTGGATTGTTTTACAAGAT G D R W A N I C R W T E D R W I ALT O D ANI CRWTEDRWV GDRW **GGAGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGAC** ATTCTTCTAAAATGGCAGCATTTTACTGAAGAACAGTGCCTTTTTAGTACATGGCTTTCA lwrs I L L K W Q H F T E E Q C L F S I L L K W Q R L T E E Q C L F S S Α **ATCCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCA**

FIG.7B-1
SUBSTITUTE SHEET

					*		*	*		*		*							
GZ	AAA	AGA	AAG	ATG	CAAT	rga/	AGAZ	ACA'	TTC	AGA	CAA	GTG	GCT"	TTA	AAG	ATC	AAA	ATG	AAATO
E		E	D		M		N				S				D				M
E			D		V	N							F						M
G	\AA/	\AG?	\AG!	ATG	CAGI	'GA	ACA	AGA:	TTC	ACA	CAA	CTG	GCT'	TTA	AAG	ATC	AAA	ATG!	AAATG
*	_					_			_		•								k *
		ים מי	ישיבי	יייין	~ \CA2	- \AA1	 РДТ(ייים.	տհոմո. 	PAA!	 'AA'	ragi	ATC:	ragi	AAA	AGAZ	AAA	AGCC	CAACC
M	3 _	S		H			S				Ī		L		K		K		T
L	s	S	L	Q	K	L	Α	V	L	K	A	ם	L	E					
TI	ATC	AAC	TCI	TC	NAAA	ACI	rggc	CCG	ΓTΤ.	[AA]	AAG	EGG 2	ATC?	rag <i>i</i>	\AA/	\GA?	LAA	\GC#	ATCC
	_							_			_					_			
200	*	: * 	A (***	*** ```	i Tarana	ای در	א מייי	א יייייייייייייייייייייייייייייייייייי	120	لمانية (ארי רים איי	-	zz CGC	יא רייז	r <i>C</i>	= \	ארא ז	CTC	AGTG
M		AAA K		S		AC1 L		Q		L			A				K	S	A WGIG
M		K		Y		L			Ď		L		T		K	N	K	s	Ÿ
AT			ACI	GTA	TTC	ACI		AĈ	\AG/	TCI		TTC	ÀAC	CACI	IGAA	GAA	TAA	GTC	AGTG
			_			_													
	*	*	* ~~~		**	** ***********************************	** ***********************************		*	Corner	·········	י מיזמי	. A.C.C	· 171-171	יככא	* ************************************	desired.	אא ממי	ירראא
_	TCA Q		M	•		M			N	F	. IGC	FO.	P	l w	D D	N N	L	T	CCAA l Q
T T		K		E	Ā				N		A	R	R C	W	Ď	N		V	lõ
		GAA	GAC								TGC	CCC	GTG	TTG	GGA				CCAA
					*			*									*		*
															CAC T			F	CCTA L
K K	L L	E E	K K		S				2	Q	A	V	T T			Q		S	
									TTC	AČA	GGC	TGI	'CÂC						ACTA
•																			
								*		*							*	*	
																			GGTA
T	Q	T	T	V	M	_				M		T	T T		_	Q	I		V
T		T GAC		V TCT	Μ ΔΔΤ	E CCA		V AGT								Q ACA			GGTA
AC	ACA.	GAC	aac	101	mı	JUN	<i></i>			1110		0110	· · · · ·		0011		···		001
	*		*									*						*	
AA	ACA	TGC	CCA	AGA														_	TGTG
K	H	A	Q	E						P					R	Q	I	T	V
K	H	A	Q TCA	E	E	L				P			K	K Gaa	R.		I Cat	T TAC	V TGTG
AA	GCA	IGC	ICA	MUM	GUA	AC1	100	acc	<i>acc</i>	ACC	100	CCA	nnn	JAA	GAG	GCA.	GAI	INC	1016
			*	*						*							*		
GA'	TTC	TGA	ACT	CAG	GAA.	AAG	GTT	GGA	TGT	CGA	TAT	AAC	TGA	ACT	TCA		TTG	GAT	TACT
D		E	L	R		R	L	D	·V	D	I	T	E	_	H	S	W	I	T
	S		I	R	K					D				_		S	W CEC	I	T
GA	TTC	TGA.	AAT.	TAG	GAA	AAG	GTT	GGA	TGT	TGA	IAI	AAC	IGA	ACT	TCA	CAG	CIG	GAI	TACT
	*				*	*		*				*	*		*	*			*
		AGA.	AGC	TGT	ATT	ACA	GAG	TTC	TGA	ATT	TGC	AGT	CTA	TCG.	AAA	AGA	AGG	CAA	CATC
R		E	A	v	L	Q	s [S	E	F	A	٧	Y	R		E	G	N	Ī
R		E	Α	V	L	Q		P	E	F	A	I	F	R	K	E	G	N	F
CG	CTC	AGA.	AGC	TGT	GTT(GCA	GAG	TCC	TGA	ATT	TGC	AAT	CTT	TCG	GAA(GGA	AGG	CAA	CTTC

FIG.7B-2

SUBSTITUTE SHEET

			**								**				*				
m/			naa:			እ አ <i>ር</i> ተባ	י מיים	N TTC	י מים	ኮአ <i>ር</i>	C	2 A C		A A C (ממר	CA	CTT	ቦሮ እ <i>ር</i>	AAAA
TC	AGA	ACT.	I GC	140/	- AAA	WG 1		310	uun.	150		יטענ יטענ	777.T.	wo.	70.UZ	10A 77	. גטר קו	יכאט	
S	D	Ţ	ΙQ	E	K	V	N	A	1	I A	K	<u> </u>		A	<u>.</u>	<u></u>	<u> </u>	R	
S	D	L	K] E	K	V	N	Α	I	E	_JR	E	K	A	E	K		R	
TC	AGA	ACT	[AA]	AG	AAA	AAGI	CA	ATG(CCA?	CAG	AGC (3AG2	LAA A	\AG(CTGA	\GA/	\GTT	CAG	AAAA
		-																	
															*	*			
											maar	7001			1000	** * *	ma		TO THE
																			TGTI
L						S	A	Q	Α	L	V	E	Q	M	A			G	V
L	O	D	Α	S	R	S	Α	Q	A	L	V	E	Q	M	V	N	E	G	V
	YCĈ I	AGZ	\ጥሮር	יראו	CAC														TGTT
~-		2.02			<i></i>														
		_	_	_		_		_									-		
		*	×	*		•		*									. .		
AA	TGC	TGA	AAC	TAT	CAC	IACA	AGC	TTC	CAGA	AAC	AAC 1	:GA/	ACAG	CCG	GTG	<u>GAC</u>	AGA	TTA	CTGC
N	A	E	s	I	R	Q	Α	S	E	Q	L	N	S	R	W	T	E	F	C
			S		K	Ιõ	Δ	S	E	ñ	T.	N	S	R	W	I	E	F	C
24	mac	<u> </u>	, D	~~~	<u> </u>	, ×	7.00	·~~	מסמי	\ A C7	\	יים א גייטי	CAG	ccc					CTGC
AA	TGU	AGA	TAC	CAI	CAA	MCA	AGC	CIL	AUA	747	WCI	.GMP	CAU		GIG	GMI	COM	WII	CIGC
	*		*			*				*			* *	:			* *	*	*
CA	ATT	'GCT	GAG	TGA	GAG	AGT	TAA	CTG	GCI	'AGA	AGTA	TCA	AAC	CAA	CAT	CAT	TAC	CTT	TTAT
	L		S	F	P	V	N	W	T.	F.	V	Ô	Т	N	I	T	T	F	Y
×	7	Ti	5	12	77	L	37	117	7	- D	v	×	1	NT	Ŧ			F	
Ž.	_L	L		_ E	K		TA	-77				_¥.		. W	~-				
CA	GTI	GCI	'AAG	TG₽	LGAG	ACT	TAA	CTG	iGC1	GGA	AGTA	TCA	GAA	CAA	CAT	CAT	CGC	111	CTAT
							*						*		*			*	*
A A	ጥሮል	CCT	מיאמי	ACA	מדים ב	CCA	262	CAT	CAC	'ΔΑ (TAC	TGC	CGA	ΑΑΑ	СТТ	GTT	GAA	AAC	CCAG
 	TCU	.00.1	.AC.	- C		- COA		M	m	~~~	~~~~ ~	, 1 O C	 E	1 M M 1	řŤ	7		T	~
N	Q	ь	Q	Ų	<u>.</u>	E	Q	M	<u> </u>	1		A	<u> </u>	7.4					
N	Q	L	Q	Q	L	E	Q	M	T	T	T	Α	E	N	W		K	LI	Q
ΑĀ	TCA	GCI	'ACA	ACA	TTA	GGA	GCA	GAT	GAC	'AAC	CTAC	TGC	TGA	aaa	CTG	GTT	'GAA	AAT(CCAA
×	± .		*									*							
				3 mc		000		200		M 2 2		CCA			א א ידי	աաշ	T A A	~~»	TGAA
S						P												D	E
Ρ	T	T	P	S	E	P	T	Α	I	K	S	Q	L	K	I		K	D	E
\overline{cc}	CAC	CAC	CCC	ATC	AGA	GCC	AAC	AGC	AAT	TAA	AAG	TCA	GTT	AAA	AAT	TTG	TAA	GGA'	TGAA
		•	**	•				•				* *					*	**	***
								~				~~.			m	ma.	<u> </u>	mam	
GT	CAA																		ACAA
V	N	R	L	S	A	L	Q	P	Q	I	E	Q	L	K	Ι	Q	S		Q
٧	N	R	L	S	G	L	0	P	Q	I	E	R	L	K	I	Q	S	I	A
ىلىتى دىلىت	~ A			ልጥሮ	AGG	ملماط	ሞፖል	ACC	тĈа	ΑΔΤ	TGA	ACG	АТТ	ΔΔΔ	ልልሞ	TCA	AAG	CATZ	AGCC
5 1	ÇE LE		001	n. (2100	101													
				_		_	_			_									_
			*	*		*	*	*		×						_		'	
																		TAC	TAAT
L	K	E	ĸ	G	0	G	P	M	F	L	D	Α	D	F	V	Α	F	T	N
- T	77	- -	W	č	~	Ğ	Ð	M	F	T.	n	Δ	ח	F	v	Δ	F	T	N
	~~		~~		. ×.			71 M	~~~			m//	202		mem.				
CI.	GAA	AUA	نAA	AGG	AUA	MUGA	MUU	CAT	GIT	CCT	GUA	TOC	MUM	CIT	TOL	GGC	CTT.	TWC	TAAA
			*	**		**	*		**		*								
מ	للململ	TAA	CCA	<u>ር</u> አጥ	بلمثنا	ፐርል'	rcc	ፐርጥ	GAG	GGC	CAA	AGA	GAA.	AGA	GCT	ACA	GAC	AATT	TTT
			H			D	- 53	77	٣	ر ا		F	17 T	म	T.	^	T	I	F
	F				F	٦	ان	V	<u>~</u>		<i>T</i> .	E-	T.	Ë	7	X	ω. T		
H	F	<u>K</u>	Q	V	F	s	<u>ט</u>	V	لعا	A	K	Ľ	K	Ľ.		Q	T	I	F
CA	TTT	TAA	GCA	AGT	CTT	TTC:	TGA	TGT	GCA	GGC	CAG	aga	GAA.	AGA	GCT.	ACA	GAC	AAT:	CTTT

FIG.7B-3

SUBSTITUTE SHEET

		*								*		*:	t			*	*	
GACA D T		'ACC P			IGCG R		Q	E	Ţ	CAAC M	S	S	I	rcac R		CGTC W	GAT I	CCAG Q
D T		P 'GCC	P 'ACC			Y CTA	Q TC	E NGGA	T AGAC			A		R CAC	T GAC			Q CCAG
GACA	CIII	.000			.000	**		•	•			•	*	*			*	
CAGT		AAG	CAA	ACI		TGI				TAG			TGA	ATA			TAA!	GGAG
Q S O S		S		L L		V				S	V V	T	E	Y	E	I	M M	E E
CAGT	CAGA	AAC	CAA		CTC	CAT	ACC	TĈA	ACT	TAG	TGI					LAAI	CAI	GGAG
*			*		*		*	*			*	*				*		* *
GAGA E R			GAA K		'ACA Q					STTC S	TTI L	K	AGA E	AGCA Q	LACA Q	N		CTTC F
Q R	L CACT	G YCGG	E GGA	L ATT	Q GCA	A GGC	L TTI	Q 'ACA	S	S TTC	L TCI	O GCA	JE Aga	Q LGCA	Q ACA	S AAG		L CCTA
CHON!	JACI	*		***	-		•		*	*			*	*	*		**	*
AACT	ATCT	'GAG																
NY	L L	S	D T	Т	A A	K K	E E	M	A S	K	K	A A	P	S		I	C S	Q R
TACT	ATCT	CAG	CAC	CAC	TGT	GAA	AGA	GAT	GTC	GAA	GAA	AGC	GCC	CTC	TGA	AAT	TAG	CCGG
AAAT	*	*		> (1941)	•m∠~ 3	7 C 7	ጵ ሮአጥ	ייייי	ccc	* *	CTP C	CAA	GNN	* A (*T	*	·	CCA	★ ሬጥጥሬ
K Y	L	1 _	E	F	E	E	I	E	G	H	W	K	K	L	S	S	Q	L
K Y AAAT	Q ATCA	S ATC	E AGA	F ATT	E TGA	E AGA	I AAT	E TGA	G .GGG	R ACG	W CTG	K GAA	K GAA				Q CCA	L GCTG
*	***		*				*	*	*				*		*			
GTGG	_	l					AGA	ACA	TAT	GAA N	TAA	ACT	TCG R		ATT		AAT N	TCAC H
V E	S H	С	Q Q	K K	L	E	E	Q	M	N	K	L	R	K	I	Q	N	H
GTTG	AGCA	TTG	TCA	AAA	GCT.	AGA	GGA	.GCA	AAT	GAA	TAA	ACT	CCG	AAA	AAT	CAG	AAT	TCAC
* ATAA	AAAC	* CTT	** ACA	GAA	ATG	GAT	GGC	TGA	AGT	TGA	TGT	TTT	* CCT	GAA	* .AGA	GGA	ATG	GCCT
IK	Ţ	L	Q	K	W	M M	A A	E	V	D D	A	F	L	K K	E E	E E	W	P P
I LQ	J T AAAC	CCT CCT	K GAA	K GAA	ATG	GAT	GGC	TGA	AGT	TGA	TGT	TTT	TCT					
	*		*	*		*	*		*		*		*					**
GCCC			TGC		AAT I	CCT L	GAA K	AAA K	ACA O	GCT L	CAA K	ACA O	ATG C	CAG R	ACT L	TTT L		TGGT G
A L	G	ם	s	E	I	L	K	K	Q	L	K	Q	С	R	L	L	V AGT	
GCCC.	rigg	GGA	TIC	AGA	WWI	ICI		nnn		.GCI		ACA	GIG	CAG	ACI		AUI	Crio I
GATA:	ITCA	* Aac	AAT	TCA	GCC	CAG	* TTT	AAA	* TAG	TGT	- Taa	TGA	AGG	TGG	GCA	GAA	GAT.	AAAG
D I	Q	T T	I	Q Q		S	L	N	S	V	N			G G	Q Q	K K	I	K K
GATA:	rtca	GAC	AAT	TČA	GCC	CAG	TCT	AAA										

FIG.7B-4

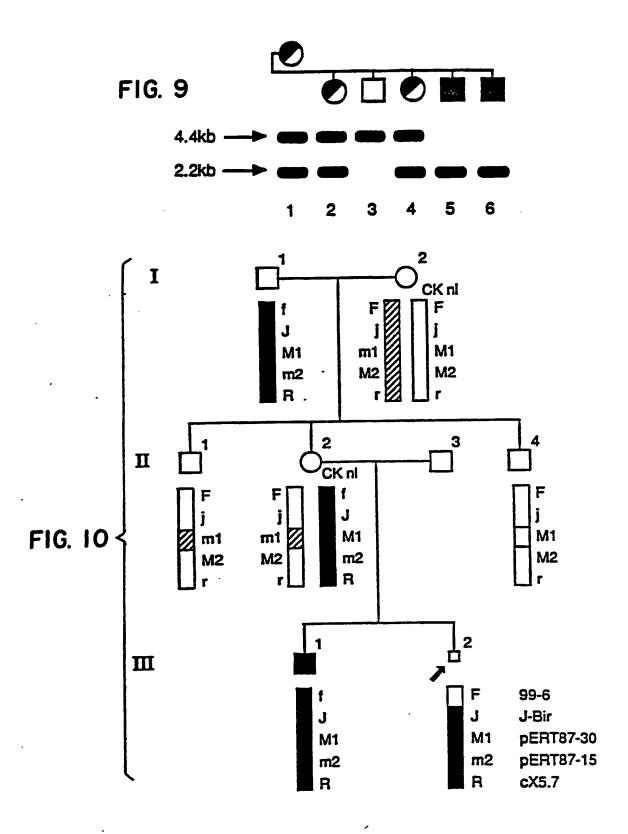
AGTGAAGCTGAACTTGAGTTTGCATCCAGACTGGAGACAGAACTTAGAGAGCTTAACACT S E A E L E F A S R L E T E L R E L N N E A E P E F A S R L E T E L K E L N AATGAAGCAGAGCCAGAGTTTGCTTCGAGACTTGAGACAGAACTCAAAGAACTTAACACT CAGTGGGATCACATATGCCGCCAGGTCTACACCAGAAAGGAAGCCTTAAAAGGCAGGTTTG Q W D H I C R Q V Y T R K E A L K A G L Q W D H M C Q Q V Y A R K E A L K G G L CAGTGGGATCACATGTGCCAACAGGTCTATGCCAGAAAGGAGGCCTTGAAGGGAGGTTTG GATAAAACCGTAAGCCTCCAAAAAGATCTATCAGAGATGCATGAGTGGATGACACAAGCT DKTVSLQKDLSEMHEWMT EKTVSLQKDLSEMHEWMT GAGAAAACTGTAAGCCTCCAGAAAGATCTATCAGAGATGCACGAATGGATGACACAAGCT GAAGAAGAATATCTAGAGAGAGATTTTGAATATAAAACTCCAGATGAATTACAGACTGCT YLERDFEYKTPDEL EEEYLERDFEYKTPDE GAAGAAGAGTATCTTGAGAGAGATTTTGAATATAAAACTCCAGATGAATTACAGAAGCA GTTGAAGAATGAAGAGGCTAAAGAAGAGGCACTACAAAAAGAAACTAAAGTGAAACTC E M K R A K E E A E M K R A K E E A EALQKETKVKL 717 GTTGAAGAGATGAAGAGAGCCCC 2153 CTTACTGAGACTGTAAATAGTGTAATAGCTCACGCTCCACCCTCAGCACAAGAGGCCTTA LTETVNSVIAHAPPSAOEAL **AAAAAGGAACTTGAAACTCTGACCACCAACTACCAATGGCTGTGCACCAGGCTGAATGGA** K K E L E T L T T N Y O W L C T R L N G **AAATGCAAAACTTTGGAAGAAGTTTGGGCATGTTGGCATGAGTTATTGTCATATTTAGAG** K C K T L E E V W A C W H E L L S Y L E **AAAGCAAACAAGTGGCTCAATGAAGTAGAATTGAAACTTAAAACCATGGAAAATGTTCCT** K A N K W L N E V E L K L K T M E N V P GCAGGACCTGAGGAAATCACTGAAGTGCTAGAATCTCTTGAAAATCTGATGCATCATTCA AGPEEITEVLESLENLMHHS GAGGAGAACCCAAATCAGATTCGTCTATTGGCACAGACTCTTACAGATGGAGGAGTCATG EENPNQIRLLAQTLTDGGVM GATGAACTGATCAATGAGGAGCTTGAGACGTTTAATTCTCGTTGGAGGGAACTACATGAA DELINEELETFNSRWRELHE GAGGCTGTGAGGAAACAAAAGTTGCTTGAACAGAGTATCCAGTCTGCCCAGGAAATTGAA EAVRKQKLLEQSIQSAQEIE AAGTCCTTGCACTTAATTCAGGAGTCGCTTGAATTC 3273 K S L H L I Q E S L E F

Nucleic Acid Homology (2153, 1885) 87.552% Amino Acid Homology (717, 621) 86.611%

FIG.7B-5

FIG. 8

MLWWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKQHIENLFSDLQDGRRLLDLLEGLTGQKLPKEKGSTR <u>VHALNNYNKALRVLQNNNVDLVNIGSTDIVDGNHKLTLGLIWNIILHWQVKNVMKNIMAGLQQTNSEKIL</u> LSWVRQSTRNYPQVNVINFTTSWSDGLALNALIHSHRPDLFDWNSVVCQQSATQRLEHAFNIARYQLGIE KLLDPEDVDTTYPDKKSILMYITSLFQVLPQQVSIEAIQEVEMLPRPPKVTKEEHFQLHHQMHYSQQITV SLAOGYERTSSPKPRFKSYAYTQAAYVTTSDPTRSPFPSQHLEAPEDKSFGSSLMESEVNLDRYQTALEE vlswllsaedtiqaqgeisndvevvkdqfhthegymmdltahqgrvgnilqlgskligtgklsedeetev QEQMNILNSRWECLRVASMEKQSNIHRVIMDLQNQKLKELNDWLTKTEERTRKMEEEPIGPDLEDLKRQV QQHKVLQE DLEQEQVRVNSLTHMVVVVDESSGDHATAALEEQLKVLGDRWANICRWTEDRWVLLQDILLK WORLTEEQCLFSAWLSEKEDAVNKIHTTGFKDQNEMLSSLQKLA VLKADLEKKKQSMGKLYSLKQDLLST LKNKSVTQKTEAWLDNFARCWDNLVQKLEKSTAQ ISQAVTTTQPSLTQTTVMETVTTVTTREQILVKHAQ EELPPPPPOKKROITVDSEIRKRLDVDITELHSWITRSEAVLQSPEFAIFRKEGNFSDLKEKVN AIEREK aekfrklqdasrsaqalveqmvnegvnadsikqaseqlnsrwiefcqllserlnwleyqnniiafynqlq OLEOMITTAENWLKIQPITPSEPTAIKSQLKICK DEVNRLSGLQPQIERLKIQSIALKEKGQGPMFLDAD FVAFTNHFKQVFSDVQAREKELQTIFDTLPPMRYQETMSAIRTWVQQSETKLSIPQLSVTDYEIMEQRLG ELQ ALQSSLQEQQSGLYYLSTTVKEMSKKAPSEISRKYQSEFEEIEGRWKKLSSQLVEHCQKLEEQMNKL RKIQNHIQTLKKWMAEVDVFLKEEWPALGDSEILKKQLKQCR LLVSDIQTIQPSLNSVNEGGQKIKNEAE PEFASRLETELKELNTQWDHMCQQVYARKEALKGGLEKTVSLQKDLSEMHEWMTQAEEEYLERDFEYKTP DELOKAVEEMK RAKEEAQQKEAKVKLLTESVNSVIAQAPPVAQEALKKELETLTTNYQWLCTRLNGKCKT LEEVWACWHELLSYLEKANKWLNEVEFKLKTTENIPGGAEEISEVLDSLE NLMRHSEDNPNQIRILAQTL TDGGVMDELINEELETFNSRWRELHEEAVRRQKLLEQSIQSAQETEKSLHLIQESLTFIDKQLAAYIADK vdaaqmpqeaqkiqsdltsheisleemkkhnqgkeaaqrvlsqidvaqkklqdvsmkfrlfqkpanfelr LOESKMILDEVKMHLPALETKSVEQEVVQSQLNHCV NLYKSLSEVKSEVEMVIKTGRQIVQKKQTENPKE ldervtalklhynelgakvterkqqlekclklsrkmrkemnvltewlaatdmeltkrsavegmpsnldse vawgk atqketekqkvhlksitevgealktvlgkketlvedklsllnsnwiavtsraeewlnllleyqkh METFDONVDHITKWIIQADTLLDESEKKKPQQKEDVLKRLK AELNDIRPKVDSTRDQAANLMANRGDHCR kivepoiseinhrfaaishriktgkasiplkeleqfnsdiqkilepleaeiqqgvnlkeedfnkdmnedn EGTVKELLQRGDNLQQRITDERKREELKIKQQLLQTKHNALKDLRSQRRKKALEISHQWYQYKRQADDLL KCLDDIEKKLASLPEPRDERKIKEIDRELQKKKEELNAVRRQAEGLSEDGAAMAVEPTQIQLSKRWREIE SKFAQFRRLNFAQIHTVREETMMVMTEDMPLEISYVPSTYLTEITHVSQALLEVEQLLNAPDLCAKDFED LFKQEESLK NIKDSLQQSSGRIDIIHSKKTAALQSATPVERVKLQEALSQLDFQWEKVNKMYKDRQGRFD RSVEKWRRFHYDIKIFNQWITEAEQFLRKTQIPENWEHAKYKWYLKELQ DGIGQRQTVVRTLNATGEEII QQSSKTDASILQEKLGSLNLRWQEVCKQLSDRKKRLEEQKNILSEFQRDLNEFVLWLEEADNIASIPLEP GKEOOLKEKLEQVK LLVEELPLRQGILKQINETGGPVLVSAPISPEEQDKLENKLKQTNLQWIKVSRALP EKQGEIEAQIKDLGQLEKKLEDLEEQLNHLLLWLSPIRNQLEIYNQPNQEGPFDVQETEIA VQAKQPDVE EIISKGOHLYKEKPATQPVKRKLE DISSEWKAVNRLLQELRAKQPDLAPGITTIGASPTQTVTLVTQPVV TKETAISKLEMPSSIMLEVPALADFNRAWTELTDWLSLLDQVIKSQRVMVGDLEDINEMIIKQK ATMODL EQRRPQLEELITAAQNIKNKTSNQEARTIITDRIERIQNQWDEVQEHLQNRRQQLNEMLKDSTQWLEAKE EAEQVIGQARAKLESWKEGPYTVD AIQKKITETK QLAKDLRQWQTNVDVANDLALKILRDYSADDTRKVF MITENINASWRSIHKRVSEREAALEETHRLLQQFPLDLEKFLAWLTEAETTANVLQDATRKERLLEDSKG VK ELMKOWODLOGETEAHTDVYHNLDENSQKILRSLEGSDDAVLLQRRLDNMNFKWSELRKKSLNIRSHL EASSDQWKRLHLSLQELLVWLQLKDDELSRQAPIGGDFPAVQKQNDVH RAFKRELKTKEPVIMSTLETV I <u>IFLTEOPLEGLEKLYQEPRELPPEERAQNVTRILRKQAEEVNTEWEKLNLHSADWQRKIDETLERLQELQ</u> EATDELDLKLRQAEVIKGSWQPVGDLLIDSLQDHLEKVK ALRGEIAPLKENVSHVNDLARQLTTLGIQLS PYNISTLEDINTRWKLIQVAVEDRVRQLHEAHRDFGPASQHFISTSVQGPWERAISPNKVPYYINHETQT TCWDHPKMT ELYGSLADINNVRFSAYRTAMKLRRIQKALCLDLLSLSAACDALDQHNLKQNDQPMDIIQI INCLTTIYDRLEQEHNNLVNVPLCVDMCLNWLLNVYDTGRTGRIRVLSFKTGIISLCKAHLEDKYRYLFK QVASSTGFCDQRRIGILLHDSIQIPRQLGEVASFGGSNIEPSVRSCFQFANNKPEIEAALFLDWMRLEPQ SMVWLPVLHRVAAAETAKHQAKCNICKECPIIGFRYRSLKHFNYDICQSCKKTGRVAKGHKMHYPMVEYC TPTTSGEDVRDFAKVLKNKFRTKRYFAKHPRMGYLPVQTVLEGDNMETPVTLINFWPVDSAPASSPQLSH DDTHSRIEHYASRIAEMENSNGSYLNDSISPNESIDDEHLLIQHYCQSLNQDSPLSQPRSPAQILISLES **EERGELERILADLEEENRNLQAEYDRLKQQHEHKGLSPLPSPPEMMPTSPQSPRDAELIAEAKLLRQHKG** RLEARMOILEDHNKQLESQLHRLRQLLEQPQAEAKVNGTTVSSPSTSLQRSDSSQPMLLRVVGSQTSDSM GEEDLLSPPQDTSTGLEEVMEQLNNSFPSSRGRNTPGKPMREDTM1



SUBSTITUTE SHEET

DYSTROPHIN →

FIG. II A

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

DYSTROPHIN → E

FIG. II B

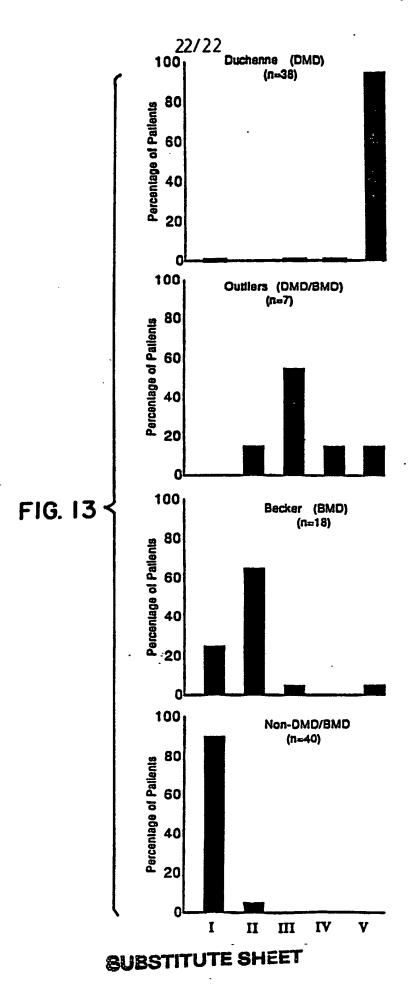
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

PCT/US88/04504 WO 89/06286

21/22

B N B N B FIG. 12

WO 89/06286 PCT/US88/04504



WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁴:
C 12 Q 1/68, C 07 H 21/00, C 07 K 15/00
C 12 P 21/02, C 12 N 15/00, G 01 N 33/68
A 61 K 45/00

(11) International Publication Number: WO 89/06286
(43) International Publication Date: 13 July 1989 (13.07.89)

(21) International Application Number: PCT/US88/04504

(22) International Filing Date: 16 December 1988 (16.12.88)

(31) Priority Application Number:

136,618

(32) Priority Date:

,ŧ

7

22 December 1987 (22.12.87)

(33) Priority Country:

US

(71) Applicant: THE CHILDREN'S MEDICAL CENTER CORPORATION [US/US]; 300 Longwood Avenue, Boston, MA 02115 (US).

(72) Inventors: KUNKEL, Louis, M.; 5 Braeburn Road, Hyde Park, MA 02136 (US). MONACO, Anthony; 51 Queenberry Street, #1, Boston, MA 02115 (US). HOFFMAN, Eric, P.; 133 Waban Street, Newton, MA 02158 (US). KOENIG, Michel; 51 Park Drive, No. 9, Boston, MA 02115 (US). (74) Agent: FREEMAN, John, W.; Fish & Richardson, One Financial Center, Suite 2500, Boston, MA 02110-2658 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), BR, CH (European patent), DE (European patent), DK, FI, FR (European patent), GB (European patent), IT (European patent), IP, LU (European patent), NL (European patent), NO, SE (European patent), SU.

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report: 10 August 1989 (10.08.89)

(54) Title: PROBES FOR AND METHODS OF DIAGNOSIS FOR MUSCULAR DYSTROPHY

(57) Abstract

The invention relates to a muscular dystrophy (MD) probe comprising a substantially purified single-stranded nucleic acid sequence capable of hybridizing to a region of DNA on a human X chromosome between the deletion break point at Xp21.3 and the translocation break point at X;11. The invention also relates to a 14kb cDNA corresponding to the complete MD gene and probes produced therefrom useful in genetic methods of diagnosis of MD. Furthermore, the invention relates to the polypeptide, dystrophin, which corresponds to the MD gene product, and antibodies thereto that are useful in a variety of methods for immunodiagnosis of MD.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	ML	Mali
AU	Australia	GA	Gabon	MR	Mauritania
BB	Barbados	GB	United Kingdom	MW	Malawi
BE	Belgium	HU	Hungary	NL	Netherlands -
BG	Bulgaria	П	Italy	NO	Norway
BJ	Benin	JP	Japan	RO	Romania
BR	Brazil	KP	Democratic People's Republic	SD	Sudan
Œ	Central African Republic		of Korea	SE	Sweden
CG	Congo	KR	Republic of Korea	SN	Senegal
Œ	Switzerland	u	Liechtenstein	SU	Soviet Union
		LK	Sri Lanka	10	Chad
CM	Cameroon				
DE	Germany, Federal Republic of	LU	Luxembourg	TG	Togo
DK	Denmark	MC	Monaco	us	United States of America
FT	Eleland	MC	Modeoccar		

INTERNATIONAL SEARCH REPORT

	International Applica	PCT/US 88/04504
L CLAS	ESIFICATION OF SUBJECT MATTER (it several clearly	
According 4	C 12 O 1/68. C 07 H 21 700 Retional Classification and	/PC
IPC:	C 12 N 15/00, G 01 N 33/68; A 61 K	15/00
II. FIELI	DS SEARCHED	
Classifica	Minimum Documentation Searched 7	
	Classification Symbols	
IPC ⁴	C 12 Q, C 12 N	·
	Documentation Searched other than Minimum Document to the Extent that such Documents are included in the Field	tation
III. DOCI	UMENTS CONSIDERED TO BE RELEVANT	
ategory *	Citation of Document, 11 with Indication, where appropriats, of the relevant p	Passages 12 Relevant to Claim No. 13
Х	Cell, vol. 50, 31 July 1987, Cell P. M. Koenig et al.: "Complete clons of the Duchenne muscular dystroph (DMD) cDNA and preliminary genome organization of the DMD gene in and affected individuals", pages 509-517 see the whole article	ing hy
Y		1-3,5-10, 35-37,40
x	Science, vol. 238, October 1987, E.P. Hoffman et al.: "Conservation of the Duchenne muscular dystroph gene in mice and humans", pages 347-350 see the whole article	on ay
Y	· -	1-3,5-10, 35-37,40
"A" docu		iblished after the international filing date of not in conflict with the application but
"E" earlie filing "L" docu which citatie "O" docu other "P" docu later t	or document but published on or after the international date ment which may throw doubts on priority claim(s) or is cited to establish the publication date of enother or other special reason (as specified) "Y" document of part cannot be considered in the priority of the international filing date but than the priority date claimed "A" document member when the priority date claimed "A" document member when the priority date claimed "The international filing date but the file of the international filing date but the priority date claimed "A" document member when the consideration of the consid	ticular relevance: the claimed invention
. CERTIF	ICATION	
	Actual Completion of the International Search Date of Mailing of this in Narch 1989 18 177.89	nternational Search Report
	Searching Authority EUROPEAN PATENT OFFICE	PLG VAN DED DITTED

Form PCT/ISA/210 (second sheet) (January 1985)

./..

Category .	Citation of Document, with indication, where appropriate, of the relevant passages	
	Dessegos	Relevant to Claim No
; ! : : !	G.J.B. Van Ommen et al.: "A physical map of 4 million bp around the Duchenne muscular dystrophy gene on the human X-chromosome", pages 499-504 see the whole article	-
Y į		1-3,5-10, 35-37,40
X	Nature, vol. 323, 16 October 1986, A.P. Monaco et al.: "Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene", pages 646-650 see the whole article	4
Y !		1-3,5-10, 35-37,40
X , X	WO, A, 88/00979 (THE CHILDREN'S HOSPITAL CENTRE CORP.) 11 February 1988 see the whole document	1-10,35-3°
,х	WO, A, 88/03572 (DAVIES et al.) 19 May 1988 see the whole document	1-10,35-37 40
X :	Nucleic Acids Research, vol. 15, no. 5, 1987, IRL Press Ltd, (Oxford, GB) T.J. Smith et al.: "Isolation of a conserved sequence deleted in Duchenne muscular dystrophy patients" pages 2167-2174 see the whole article	1-10,35-37 40

,		•
:		
•		:
:		

Form PCT ISA 210 (extra sheet) (Jenuary 1985)

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET	
V. OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE	
This international search report has not been established in respect of certain claims under Article 17(2) (a) for the	he following reasons:
1. Claim numbers because they relate to subject matter not required to be searched by this Authority	
•	
• •	
2. Claim numbers	the prescribed require-
ments to such an extent that no meaningful international search can be carried out, specifically:	
3. Claim numbers because they are dependent claims and are not drafted in accordance with the second PCT Rule 6.4(a).	and third sentences of
VI OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING 2	
This international Searching Authority found multiple inventions in this international application as follows: Please refer to Form PCT/ISA 206 dated 17th April:	1989
As all required additional search fees were timely paid by the applicant, this international search report covers of the International application.	s all searchable claims
2. As only some of the required additional search fees were timely paid by the applicant, this international sea those claims of the international application for which fees were paid, specifically claims:	rch report covers only
3. No required additional search fees were timely paid by the applicant. Consequently, this international search the invention first mentioned in the claims; it is covered by claim numbers: 1-3 and partially 4-10;	
An all searchable claims could be searched without effort justifying an additional fee, the International Search	Ť
Invite payment of any additional fee. Remark on Protest	-
The additional search fees were accompanied by applicant's protest.	2
No protest accompanied the payment of additional search fees.	

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO.

US 8804504 SA 26051

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on 04/07/89

The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Pate: mer	nt family nber(s)	Publication date
WO-A- 8800979	11-02-88	EP-A-	0275298	27-07-88
WO-A- 8803572	19-05-88	None		
•				
			t	•